Three-Dimensional Mapping of Water Surface Waves using Air-Coupled Sonar

Aidan Fitzpatrick, Ajay Singhvi, Jordy Mukania, Brion Ye, Eastan Giebler, and Amin Arbabian
Department of Electrical Engineering, Stanford University, Stanford, CA, USA

Abstract—More than 70 percent of the Earth’s surface is covered by the ocean and other bodies of water, making them abundant sources of valuable information. In particular, the water surface can be sensed and mapped to extract key parameters related to water dynamics. These parameters have far-reaching implications spanning industrial, environmental, energy, navigation, and various other applications. Existing approaches for measuring the water surface each have their own respective limitations, thus leaving high-resolution spatial and temporal water surface wave mapping an open challenge in the research community. This work proposes a non-contact, acoustic surface mapping system that uses air-coupled ultrasonic transducers to capture three-dimensional spatial maps of the water surface via active sonar imaging. Within, we provide a holistic overview of the system’s design parameters before narrowing in on details related to the hardware, ultrasonic transmit waveforms, and signal processing pipeline. Through verification in simulation, the proposed air-coupled sonar system demonstrates high-fidelity surface mapping with millimeter-scale accuracy and spatial resolution from standoff distances up to several meters above the water.

Index Terms—air-coupled, acoustic, non-contact, reconstruction, sonar, ultrasound, water surface mapping, waves

I. INTRODUCTION

Accurate information about water dynamics is of great importance for a number of applications including environmental monitoring [1], marine navigation [2], infrastructure engineering [3], as well as climate and other scientific research [4]. Within these domains, the ability to capture information about surface waves enables more accurate weather prediction [5], allows for real-time navigation planning through safer and more efficient routes [2], provides understanding of hydraulic forces to ensure resilient infrastructure design [3], and aids in modeling hydrological processes for activities including flood prediction and water resource management [6].

In addition to the above applications, emerging technologies for airborne imaging [7]–[9] and communication [10]–[12] across the air-water boundary require accurate spatial information about the water surface. Due to the vastly different medium properties of water and air, both electromagnetic and acoustic waves incur refraction and distortion effects that must be corrected to maintain high-fidelity imaging and communication links through dynamic water surfaces [11], [13]. As a result, these technologies are in search of a robust means for high-resolution, three-dimensional (3-D) spatial mapping of water surface waves.

There are a number of sensors, ranging from experimental approaches to well-established commercial products, that can measure various parameters of the water surface. In-situ sensors, such as wave gauges [14], wave buoys [15], and acoustic doppler current profilers (ADCPs) [16], provide highly accurate, temporally resolved but localized measurements of parameters such as wave velocity, height, direction, and spectra. While in-situ sensors are the gold standard for accuracy and are often used to ground truth emerging approaches [17], they inherently lack spatial coverage and spatially resolved wave features.

On the other hand, non-contact measurements using cameras [18], radar [19], lidar [20], or ultrasonics [11], can provide both spatially and temporally resolved wave fields from which parameters of interest can be extracted [21]. To date, non-contact sensors have been limited to either high-resolution, lab-based measurements that are not practical in real-world conditions [22], [23] or field-based implementations that sacrifice resolution and data quality for spatial coverage and robustness [19], [20]. As such, these non-contact approaches complement the capabilities of in-situ tools, but there is still a dearth of solutions that can capture high-resolution spatial and temporal surface elevation measurements, particularly of small-scale wave patterns in real-world conditions.

To address this critical sensing gap, this work presents an acoustic surface mapping system that leverages sonar principles to provide 3-D surface measurements \((x, y, z)\) with millimeter-scale accuracy and spatial resolution from standoffs up to several meters above the water. We envision that such a non-contact surface mapping solution could be deployed from platforms that are either stationary (e.g., bridges) or dynamic (e.g., drones) across a range of environments, including inland water bodies.
rivers and lakes, coastal waters, and open oceans.

In the remainder of this paper, we articulate the key components of the system concept, discuss the underlying principles to explore the corresponding design tradeoffs, present our methodology for data acquisition and processing, and lastly validate the concept through a number of end-to-end simulations. Future work will implement and experimentally demonstrate the proposed system.

II. SYSTEM OVERVIEW

A. System Concept

Fig. 1 illustrates the system concept which uses an air-coupled ultrasonic transducer array to capture 3-D spatial maps of the water surface via active sonar imaging. First, ultrasound transmitters (TX) insonify the water surface from a standoff in air. Due to the high mechanical contrast of water and air, over 99% of the acoustic signal is reflected, and the resulting echoes are detected by ultrasound receivers (RX). Interfacing electronics digitize the captured signals and a digital signal processing pipeline outputs a 3-D reconstruction of the instantaneous water surface. With sufficient frame rate, such a system can extract key parameters including wave velocity, height, direction, and spectra over a high-resolution 3-D wave field. Further details on the specific hardware architecture, transmit waveform, and processing pipeline are provided in the subsequent sections.

B. System Specifications

The proposed system offers various design knobs that can be tuned to meet application-specific requirements and measurement objectives. Important application-dependent operational specifications include the 1) standoff in air, 2) spatial resolution, 3) depth resolution, and 4) frame rate (or temporal resolution). In this section, we discuss the key system design trade-offs before our specific design decisions are presented in detail in Section II-C and Section II-D.

1) Standoff Height: There is a strong analogy that can be drawn between the proposed air-coupled sonar for mapping the water surface and conventional in-water sonar used for mapping bathymetry. Specifically, the standoff height in air is analogous to the sounding depth in water. It is well-understood that the maximum sounding depth of a bathymetric sonar is inversely related with frequency, that is, as frequency increases, the penetration depth decreases due to a higher degree of acoustic attenuation.

A similar, yet more severe trade-off exists for an air-coupled sonar. For example, at 100 kHz acoustic frequency, the acoustic attenuation in water is typically on the order of 0.05 dB/m whereas in air, the attenuation is over 50 times greater at approximately 3 dB/m [24], [25]. As such, acoustic attenuation, and thus system signal-to-noise ratio (SNR), is the primary factor which limits the maximum standoff. Fig 2(a) depicts the expected acoustic attenuation in dB as a function of acoustic frequency and standoff height.

2) Spatial Resolution: Despite incurring additional acoustic attenuation, bathymetric sonars often push to use higher frequencies to obtain higher spatial resolution when not SNR-limited. The spatial resolution of a monostatic imaging system is:

$$\Delta x = \frac{H\lambda}{L_x}, \quad \Delta y = \frac{H\lambda}{L_y}$$  \hspace{1cm} (1)

where $\Delta x$ and $\Delta y$ are the spatial resolutions in the $(x,y)$-directions, $H$ is the height of the system, $L_x$ and $L_y$ are the respective lengths of the system aperture in the $(x,y)$-directions, and where $\lambda$ is the acoustic wavelength and is a function of the speed-of-sound in the medium ($c$) and the acoustic frequency ($f$):

$$\lambda = \frac{c}{f}.$$  \hspace{1cm} (2)

For simplicity, angle dependence is neglected in (1).

It can be seen through (1) and (2) that the spatial resolution improves (i.e. smaller $\Delta x$ and $\Delta y$) as the acoustic frequency increases and as the system aperture size increases; Fig. 2(b) illustrates this relationship.

Through Fig. 2(a)-(b), it is clear that the acoustic frequency is an important design knob that must be tuned to effectively balance attenuation and resolution for a given application. For example, if the application requires centimeter-level spatial resolution to resolve small wind-driven waves, it is likely that the system will need to reduce standoff height to operate at a higher acoustic frequency while mitigating severe attenuation. Conversely, if the application is interested in longer gravity
waves, spatial resolution could be sacrificed to operate at lower acoustic frequencies and thus higher standoffs.

In addition to the acoustic frequency, Fig. 2(b) shows that spatial resolution is also driven by the size of the system aperture. Often, system aperture size is constrained by practical considerations such as the cost and complexity of designing and building large transducer arrays and will be discussed in more detail in Section II-C.

3) Depth Resolution: The depth resolution ($\Delta z$) of the system is highly related to the accuracy with which the water surface elevation in the $z$-direction can be measured and is a function of the system’s bandwidth ($B$):

$$\Delta z = \frac{c}{2B}. \quad (3)$$

Fig. 2(c) plots the depth resolution of the system as a function of bandwidth. Again, the angle dependence is neglected in (3) for simplicity.

The bandwidth of the system is determined primarily by two factors: the available bandwidth of the hardware (transducers and electronics) as well as the bandwidth of the transmit signal. Within the constraints of the hardware, the transmit waveform can be designed to achieve a desired bandwidth and thus depth resolution; this will be discussed in more detail in Section II-D.

4) Frame Rate: With a single captured frame, the proposed system can reconstruct a 3-D spatial map of the surface wave elevation; however, to extract additional information about the wave dynamics such as the wave velocity and direction, a number of frames must be captured with sufficient frame rate, or temporal resolution. Per Nyquist sampling theorem, unambiguous temporal sampling is achieved with a frame rate that is twice the maximum temporal frequency ($f_w$) of the wave field [26]:

$$f_w = \frac{\sqrt{k_w g}}{2\pi}, \quad k_w = \frac{2\pi}{L_w} \quad (4)$$

where $g$ is the gravitational acceleration constant $g = 9.81$ m/s$^2$ and $L_w$ is the wavelength of the wave component. From (4) it can be noted that smaller wavelengths have higher temporal frequencies. Therefore, to satisfy Nyquist sampling, the frame rate of the system must satisfy:

$$F_r > 2f_w \approx \frac{2.5}{\sqrt{\lambda_{min}}}. \quad (5)$$

where $F_r$ is the frame rate of the system in frames per second and $\lambda_{min}$ is the minimum wavelength desired to resolve.

C. Hardware Architecture

As discussed above, the design of the transmitter and receiver arrays is motivated by the desired resolution while balancing considerations including the form factor, cost, and complexity of the system. While increasing the size of the aperture, i.e. $L_x$ and $L_y$, enhances the spatial resolution of the system, larger size typically equates to more elements in the receiver array. With an increase in the number of elements, there is a corresponding rise in implementation cost and complexity due to the need to support additional peripheral hardware, handle larger data volumes, and provide greater processing power.

To increase the spatial resolution of the system without increasing the physical size of the receiver aperture, we propose leveraging a multistatic configuration [27] of transmitters and receivers as depicted in Fig. 3(a). In a multistatic imaging system employing multiple transmitters, the spatial resolution is no longer limited by the physical size of the real aperture, but instead the size of the effective aperture which can be found as the spatial convolution of the transmitter and receiver apertures [27]. By strategically placing the locations of the transmitters, we achieve an effective aperture that is $2L_x \times 2L_y$ in size as shown in Fig. 3(b). Correspondingly, the multistatic spatial resolutions for this design can be found as:

$$\Delta x = \frac{H\lambda}{2L_x}, \quad \Delta y = \frac{H\lambda}{2L_y}, \quad (6)$$

where $L_x$ and $L_y$ are the respective lengths of the real aperture in the $(x, y)$-directions.

A comparison of the spatial resolution achieved by a 25 cm array of receivers using one transmitter versus four transmitters in a multistatic configuration is shown via point spread function [28] simulations in Fig. 3(c)-(d); as expected, the multistatic array demonstrates $2 \times$ better spatial resolution.
The illustrated point spread functions are simulated for a 70 kHz acoustic frequency, which is a design decision that will be discussed in detail in Section II-D. Considering this selected frequency, the receiver aperture is designed such that \( L_x = L_y = 25 \text{ cm} \) to achieve sub-centimeter spatial resolution at a 1 m standoff.

Despite leveraging a multistatic array, it should be noted that a 25 cm \( \times \) 25 cm receiver array requires approximately 10,000 elements to maintain \( \lambda/2 \) element pitch at 70 kHz acoustic frequency, which is prohibitive in practice. As a result, our future implementation plans to combine this multistatic configuration with a sparse receiver array known to suppress grating lobes while maintaining equivalent spatial resolution using far fewer elements [29]–[31].

**D. Transmit Waveform**

The transmit waveform contains several degrees-of-freedom including the signal modulation scheme, frequency, and bandwidth that must also be tuned to meet the system requirements.

1) **Modulation Scheme**: One challenge of the multistatic hardware architecture proposed in Section II-C is the requirement that the transmit signals be orthogonal in some domain such that the signals can be delineated at the receivers. Generally, this can be achieved through time-division, frequency-division, or code-division multiplexing [32]. For our application of mapping dynamic water waves, the transmitters cannot be time division multiplexed, or fired sequentially, as the wave pattern seen by each transmitter will vary. Likewise, frequency division multiplexing requires allocating transmit signals in different frequency bands, thus reducing the bandwidth of each and therefore sacrificing depth resolution of the system.

We propose using code division multiplexing (CDM) to enable simultaneous firing of the transmitters over the entirety of the available bandwidth. For CDM, each of the transmit waveforms is phase modulated with sequences of pseudorandom codes that have excellent cross-correlation properties [33]. After simultaneous detection at the receivers, signals can be delineated through matched filtering, or cross-correlation, with each of the transmit waveforms [34]. Processing can then be performed independently as if the transmitters were fired sequentially.

In addition to phase modulation enabling simultaneous transmitting via CDM, it is also advantageous for increasing SNR. The SNR of the matched filter output is given by [35]:

\[
SNR = \frac{2P_x \tau_p}{N_0}, \tag{7}
\]

where \( P_x \) is the power in the signal, \( N_0/2 \) is the white noise power spectral density of the receiver system, and \( \tau_p \) is the pulse width of the sequence.

Since (7) is linearly proportional to the pulse width \( \tau_p \), SNR can be gained with longer transmit waveforms. In general, longer CDM sequences also enhance the cross-correlation properties and thus separability of the transmitters. That said, the transmit waveforms cannot be infinitely long and are inherently constrained by a short interval during which the water surface waves can be assumed to be quasi-static. Transmit waveforms that are too long will result in motion blur that will effectively reduce the resolution of the reconstructed surface.

To balance motion blur with the benefits of increased separability and SNR gained through integration time, we assume a quasi-static interval and thus maximum sequence length of 5 milliseconds. Due to the long temporal periods of water waves, which are on the order of 100’s of milliseconds to minutes or more [26], negligible motion blur is expected within the assumed 5 millisecond quasi-static interval.

The CDM sequences of length \( \tau_p = 5 \text{ ms} \) are comprised of \( N_s \) symbols of length \( \tau_s \) where each symbol is phase modulated with a pseudorandom binary code. The center frequency and bandwidth of the CDM sequence is controlled by the carrier frequency and bandwidth of a single symbol.

2) **Frequency & Bandwidth**: To select the carrier frequency and bandwidth of the symbol, we consider the measurement objective for our application: millimeter-scale depth accuracy and spatial resolution such that the system can resolve small-scale surface waves from standoffs up to several meters.

Capillary waves, or waves with wavelengths less than approximately 2 cm, carry little energy and have negligible amplitudes that are typically on the order of a few millimeters or less [36]. As such, our design goal is to have sufficient spatial resolution to resolve wavelengths greater than 2 cm from a 1 m standoff. At a 5 m standoff, this equates to having the capability to map wavelengths greater than 10 cm, which even still are expected to have amplitudes on the order of millimeters.

To achieve this design specification, we must spatially Nyquist sample the water surface such that we have 1 cm spatial resolution or better at a 1 m standoff or equivalently a 5 cm spatial resolution or better at a 5 m standoff. Balancing aperture size, attenuation, and resolution as described in Section II-B, we land on a 70 kHz acoustic frequency and a real aperture size of 25 cm as discussed in Section II-C. As was shown in Fig. 3(d), this design achieves a sub-centimeter spatial resolution at a 1 m standoff.

The bandwidth of the signal, which dictates the depth resolution, can be found as:

\[
B \approx \frac{1}{\tau_s} = \frac{f}{N_c}, \tag{8}
\]

where \( \tau_s \) is the symbol’s pulse width comprised of \( N_c \) cycles at a frequency \( f \).

Referring to (3) or Fig. 2(c), 35 kHz of bandwidth results in a depth resolution of approximately 5 mm; thereby enabling surface mapping with millimeter-scale depth accuracy. By using \( N_c = 2 \) cycles at a center frequency of 70 kHz, the symbol can achieve 35 kHz of bandwidth.

Since future work plans to implement the design proposed herein, it should be noted that commercial availability of ultrasound transmitters and receivers was also considered in the selection of a 70 kHz center frequency with 35 kHz bandwidth.
3) Summary of Waveform Design: The transmit waveform design is summarized in Fig. 4. In Fig. 4(a)-(b), each symbol consists of 2 cycles at a 70 kHz center frequency. In Fig. 4(c), the first few symbols of one of the transmit waveforms is shown to illustrate the sequence of phase modulated symbols. In total, each of the waveforms (one for each transmitter) consists of $N_t = 175$ symbols such that $T_p = 5$ ms. Such a waveform has a theoretical 45 dB SNR gain with respect to a waveform consisting of only a single symbol. This gain is especially important for combating the expected 5 dB of acoustic attenuation per meter standoff when operating at 70 kHz.

Fig. 4(d)-(e) demonstrate the separability of the transmit waveforms. By evaluating the cross-correlation matrix, we find greater than 15 dB separability between any two transmit sequences. Furthermore, Fig. 4(d) illustrates that the matched filter output achieves the designed 5 mm depth resolution.

III. SIMULATION ANALYSIS

A. Simulation Setup

In this section, we simulate the proposed system concept using a custom developed acoustic forward simulator that exploits a technique known as the Hybrid Angular Spectrum Method [37]. The simulation setup assumes the hardware architecture presented in Fig. 3(a) where the transmitters are fired simultaneously using the code division multiplexed transmit waveforms designed in Section II-D. In summary, the 4 transmitters each fire a 5 ms long sequence consisting of 175 symbols where each symbol is 2 cycles at a 70 kHz center frequency. The receiver array is 25 cm × 25 cm with an element pitch of 2 mm in the (x, y)-directions. The standoff height of the system is set to 3 m for all simulations presented herein.

To define the water surface, wave statistics are drawn from the JONSWAP spectrum [38] to first generate surface waves over a 50 m × 50 m domain. Subsequently, a 1 m × 1 m patch of the larger domain is randomly selected to define the water surface over the simulator’s computational grid. The grid size is limited to 1 m × 1 m by computational constraints. The wave spectra is varied across simulations by changing parameters of the JONSWAP spectrum including the significant wave height and peak wave frequency. Wave dynamics are not considered in the simulation as it is expected that the surface is quasi-static during the 5 millisecond integration time.

B. Processing Pipeline

To reconstruct the water surface waves from the captured acoustic signals, the processing pipeline outlined in Fig. 5 is employed. The output of the simulator is an $(N_r, N_t)$ array containing the time-domain acoustic signals of length $N_t$ captured by each of the $N_r$ receivers.

First, the raw signals are matched filtered via cross-correlation with each of the 4 transmit waveforms to delineate the signals corresponding to each of the transmitters. The resultant $(N_r, N_t, 4)$ array of matched filtered signals is then passed to a frequency-domain backprojection algorithm [39] which outputs a reconstructed image defined over the discretized computational grid of size $(N_x, N_y, N_z)$. A dynamic thresholding algorithm is then run over the 3-D reconstructed image to generate a point cloud of $N_p$ surface points. The $N_p$ surface points, each consisting of a coordinate in $(x, y, z)$, are passed to a surface fitting module which performs outlier rejection and interpolation to generate the final 3-D reconstructed surface.

C. Simulation Results

To demonstrate the efficacy of the proposed concept in various water conditions (or sea states), we simulate the ground truth wave profiles illustrated in Fig. 6(a)-(d). Fig. 6(e)-(h) presents the corresponding reconstructed surfaces after employing the processing pipeline outlined above. To quantify the performance, Fig. 6(i)-(l) plots the absolute error of the reconstruction with respect to the ground truth. The Mean Absolute Error (MAE) of the reconstruction is also presented for each example demonstrating the desired millimeter-scale accuracy on average.
Fig. 5. Overview of the processing pipeline used to reconstruct the 3-D surface. First, the simulated time-domain signals are cross-correlated (or matched filtered) with the transmit waveforms. An image is formed through a frequency-domain backprojection algorithm on the filtered signals. Dynamic thresholding is employed to extract a point cloud from the image. The final 3-D surface is formed through a surface fitting (or interpolation) of the 3-D point cloud.

Fig. 6. (a)-(d) Ground truth water surface waves input to the simulator, (e)-(h) Reconstructed surface waves after employing the processing pipeline on the simulated output data, (i)-(l) Spatial map of the absolute error between the reconstructed and ground truth surface waves; Mean Absolute Error (MAE) is noted for each.
IV. CONCLUSION

In this work, we introduced an air-coupled sonar system for three-dimensional mapping of water surface waves. The system concept and trade-off space were presented in detail to elucidate how such a system can be designed to meet various measurement objectives. We proposed an example system design aiming to achieve millimeter-scale surface mapping accuracy from a standoff height of several meters in air. Simulations of the proposed design demonstrated high-fidelity results that met the design goals across a range of wave conditions.

Future work will implement and experimentally validate the system while also analyzing the impacts of second-order effects including sea spray, wave breaking, and other real-world conditions.

REFERENCES


