Abstract

Nowadays all kinds of signals can be transmitted by powerful smartphones, which enables us to achieve more functionalities on mobile devices like tracking and imaging applications. We aim to build an acoustic imaging system on mobile device based on Synthetic Aperture Radar (SAR) algorithm. SAR or inverse SAR requires continuous collection of signal data. However, in practice due to motion error and interrupted environment, issue of sparse aperture data arises when some periods of aperture are invalid or missed. At the same time, imaging speed is also important for mobile application, which is usually slowed down by large amount of acquired data with Nyquist criterion and high computation cost of Fourier reconstruction process. Fortunately, many real world signals including audio image signal are sparse. Sparse recovery allows signal acquisition with fewer measurements and image recovery in faster time. In our project, we want to leverage compressed sensing methods to recover a perfect image with incomplete aperture data. We first propose a new method for sparse aperture imaging on mobile device and then recover signal from sparse aperture in azimuth bin. As for time efficiency, we speed up Fourier Transform (FT) image reconstruction process by sparse FFT algorithm and use only a subset of the input data required by the FFT. Simulation results proved the feasibility and superiority of our approach.
1 Introduction

Nowadays smartphones are more and more powerful that all kinds of signals can be transmitted and utilized, like sensors, light, audio, Radio Frequency (RF) and even vibration signals. With such much available data, in recent years more and more functionalities have been researched and developed on mobile devices, like positioning, tracking and imaging applications. Among all types of signal sources, acoustic signal has a dominant position. Comparing to RF signals, acoustic signal has longer wavelength so that it can achieve higher tracking accuracy and finer imaging resolution. What’s more, acoustic imaging works well in the dark environment covering the shortage of traditional light-based imaging approach. Even more amazing is that acoustic signals can penetrate some materials such as clothes and cardboards, which makes it possible to generate images of an object under clothes or in boxes. Such capability facilitates various applications, especially when such imaging system can be implemented in a mobile platform.

In order to implement a practical acoustic imaging system on mobile device, we choose to use Synthetic Aperture Radar (SAR) imaging strategy. Other imaging approaches like Computed Tomography (CT), Holography and Magnetic Resonance Imaging (MRI) can also be improved by compressed sensing in data acquisition and recovery process. Most famously, Michael et al. [LDP07] developed a framework for using CS in MRI which showed that the sparsity of MR images can be exploited to significantly reduce scan time, or alternatively, improve the resolution of MR imagery.

However, problems would arise if we want to implement SAR system onto a mobile device. First, there would rise sparse aperture data with interrupted measurements or invalid collections. Jun et al. in [JMS09] said that in actual applications, issue of sparse aperture data arises when uninterrupted measurements are impossible or the collection data during some periods are not valid. For example, ISAR transmits narrow-band and wide-band pulse signal alternately, the former determine the reference range which is used in dechirping receiver and the latter is used to imaging. Sparse aperture appears when the periods between the two pulses are not equal or some data missed. As the corrupted data may reduce the resolution in azimuth, we usually fill the gapped data using prediction and interpolation. Theoretical analysis and experience indicate that the longer the signal interpolation length is, the larger the interpolation error is. If the vacant aperture becomes bigger, the performance of above method will reduce rapidly. Another problem is that Fourier image reconstruction process has high computation cost when the dimension of Fourier Matrix is high. Due to limited azimuth accuracy and computation capability of mobile system, as well as high resolution and quick response time requirement, implementing acoustic imaging on mobile device seems more challenging.

According to the recently developed mathematical theory of compressed-sensing, images with a sparse representation can be recovered from randomly undersampled $k$-space data, provided an appropriate nonlinear recovery scheme is used. Related works [LDP07, LDSP08] have also shown that the sparsity of MR/CT images can be exploited to significantly reduce scan time, or alternatively, improve the resolution of MR/CT imagery. To resolve the issue of minimization of non-zeros elements, traditional approaches such as [TG07] orthogonal matching pursuit (OMP) and basis pursuit (BP) may be used. Non-adaptive means should perform better than FFT imaging, especially in azimuth focusing of SAR/ISAR. A nonlinear thresholding scheme can recover the sparse coefficients, effectively recovering the image itself. From the theory of compressive sensing (CS), we know that the exact recovery of an unknown sparse signal can be achieved from limited measurements by solving a sparsity-constrained optimization problem. For inverse synthetic aperture radar (ISAR) imaging [JMS09], the backscattering field of a target is usually composed of contributions by a very limited amount of strong scattering centers, the number of which is much smaller than that of pixels in the image plane. With such sparsity assumptions, we can also apply sparse FFT
algorithms proposed by [HIKP12b] to speed up the image reconstruction speed.

In our research project, we want to apply Sparse Fourier Transform and Compressed Sensing techs onto acoustic imaging processing, to reduce scanning time and sample measurements, and improve the imaging resolution. By leveraging the sparsity of received signal, it’s promising to reduce the amount of data and acquisition time and improve the anti-jamming capability. To be specifically, we first propose a novel mobile framework for ISAR imaging is proposed through Frequency-Modulated Continuous-Wave (FMCW). By using such a framework, the measurements, only at some portions of frequency subbands, are used to reconstruct full-resolution images by exploiting sparsity. Compressed sensing enables the perfect recovery of signals and data from what appear to highly sub-Nyquist-rate samples, so we next recover signal from sparse aperture in azimuth bin and propose a new method for sparse aperture imaging on mobile device. Finally we speed up Fourier Transform image reconstruction process and use only a subset of the input data required by the FFT. Simulation results show the advantages of our approach.

2 Related Work

2.1 Synthetic Aperture Radar System

Synthetic-Aperture-Radar (SAR) techniques [MPIY+ 13, Sou99] are widely adopted for Earth remote sensing, using radio-frequency (RF) waves with the frequency from 250 MHz to 40 GHz. In SAR radar, the receiver equipped with a single antenna moves along a predefined trajectory (e.g., a straight line and a circle), and collects the RF signals reflected by the scene. Figure 1 shows the illustration of the SAR spotlight operation mode. SAR is a form of radar mounted on a moving platform that is used to create images of objects. To create a SAR image, successive pulses of radio waves are transmitted to illuminate a target scene, and the echo of each signal is received and recorded.

![SAR system model](image)

Figure 1: SAR system model

A common technique for many radar systems is to “chirp” the signal. In a chirped radar, the pulse is allowed to be much longer. A longer pulse allows more energy to be emitted, and hence
received, but usually hinders range resolution. But in a chirped radar, this longer pulse also has a frequency shift during the pulse (hence the chirp or frequency shift). When the chirped signal is returned, it must be correlated with the sent pulse. Classically, in analog systems, it is passed to a dispersive delay line that has the property of varying velocity of propagation based on frequency. This technique "compresses" the pulse in time – thus having the effect of a much shorter pulse while having the benefit of longer pulse length. Newer systems use digital pulse correlation to find the pulse return in the signal.

Figure 2: SAR components

Figure 2 indicates the basic components of a SAR system and the data transfers among them. The digital recorder is not necessary in real-time operation. Functions of the pre-processor may include real-to-complex data conversion and presumming in azimuth. Auxiliary data processing involves conversion of motion sensor data to a position history of the APC. This processed motion data provides parameters to the IFP for motion compensation.

Based on the phase information of the received signals, various imaging algorithms such as Phase Format Algorithms (PFA), Range Mitigation Algorithm (RMA), and Chirp Scaling Algorithm (CSA) [Sou99], can be applied to generate images of the scene. These algorithms require the knowledge about the distance between the receiver and the scene center. We implement PFA algorithm on our system and we will talk it in detail in section 3.

2.2 Compressed Sensing in Imaging

Compressed sensing (CS) aims to reconstruct signals and images from significantly fewer measurements than were traditionally thought necessary. Works have been done on Compressed Sensing applied on other kinds of imaging approach. Magnetic Resonance Imaging (MRI) is an essential medical imaging tool with an inherently slow data acquisition process. In [LDSP08], the authors applied CS to MRI which offered potentially significant scan time reductions. Practical incoherent under-sampling schemes are developed and analyzed by means of their aliasing interference. Incoherence is introduced by pseudo-random variable-density under-sampling of phase-encodes. The reconstruction is performed by minimizing the $l – 1$ norm of a transformed image, subject to data fidelity constraints. Examples demonstrate improved spatial resolution and accelerated acquisition for multi-slice fast spin echo brain imaging and 3D contrast enhanced angiography.
Generally speaking, we assume that a discrete signal $x$ with length $N$ is $k$-sparse if at most $k$ nonzero elements. Define an $N \times N$ basis matrix $\Phi := \{\phi_1, \phi_2, \ldots, \phi_N\}$, then we can express any signal $x \in \mathbb{R}^N$ as:

$$x = \Phi s = \sum_{i=1}^{N} s_i \phi_i$$

where $s$ is an $N \times 1$ column vector showing the weighting coefficients. In order to get a measurement matrix $\Theta$ with $M \times N$ dimensions, we need to find an $M \times N$ basis matrix satisfying $\Theta := \Phi \Psi$. Then we can get measurement $y$ as:

$$y = \Phi x = \Phi \Psi s = \Theta s$$

After that, we need to use some reconstruction methods including greed algorithms such as orthogonal matching pursuit (OMP) or Basis pursuit (BP) and solving the convex problem:

$$\min \|s'\| \text{ such that } \Theta s' = y$$

### 2.3 Sparse Fast Fourier Transform

The discrete Fourier transform (DFT) is one of the most important and widely used computational tasks. Its applications are broad and include signal processing, communications, and audio/image/video compression. Hence, fast algorithms for DFT are highly valuable. Currently, the fastest such algorithm is the Fast Fourier Transform (FFT), which computes the DFT of an $n$-dimensional signal in $O(n \log n)$ time. The existence of DFT algorithms faster than FFT is one of the central questions in the theory of algorithms.

A general algorithm for computing the exact DFT must take time at least proportional to its output size, i.e., $\Omega(n)$. However, most of the Fourier coefficients of a image/audio signal are small or equal to zero, i.e., the output of the DFT is (approximately) sparse.

We would like to apply technique of sparse FFT [HIKP12a] to speed up the Fourier transform process as well as remove the noise of our signal.

At a high level, sparse Fourier algorithms work by binning the Fourier coefficients into a small number of bins. Since the signal is sparse in the frequency domain, each bin is likely to have only one large coefficient, which can then be located (to find its position) and estimated (to find its value). We use some good property filter to manipulate signal in time domain which can be seen as a binning in frequency domain. In order to get a sub-linear time algorithm, our filter in time domain should have sub-linear support size while its approximate corresponding frequency domain filter should be like a box filter in order to separate the frequency domain into small bins.

As a experiment, we would also like to find the appropriate ‘sparsity’ parameter to reconstruct an acoustic image with a good quality and high transforming speed as well.

### 3 Problem Statement and Formulation

Synthetic Aperture Radar/Inverse SAR image the targets on the assumption that the full aperture and complete radar echo is obtained. We first assume the full aperture data model is $x$ as:

$$x(n) = \sum_{i=1}^{k} a_i e^{j\omega_i n} + w(n)$$

where $n = 1, 2, \ldots, N-1$ is the number of samples, $j = \sqrt{-1}$, $k$ denotes the total number of signals, $a_i$ and $\omega_i$ denote the amplitude and frequency of the $i$th signal, and $w(n)$ is the white Gaussian noise.
But in actual situations, continuous measurements are not possible or the period between the measurements are not equal, especially in ScanSAR or ISAR systems. In our project, we assume that the data are completely compensated in range cell, but the period of dwell time is not equal and some echoes are missed as shown in the first figure in Figure 4.

Assume there are $M$ invalid subapertures and the lengths of subapertures are different. Denote the vector of the $m$th subaperture by $x_m = [x_m(0), x_m(1), \ldots, x_m(L_m)]^T$, $m = 1, 2, \ldots, M$, which are valid and the data between these subaperures are missing. $L_m$ denotes the length of $m$th subaperture data. For simplicity, the missing data are set to zeros, and then a fast Fourier transform is performed to process the zeros-filled data, which will lead high side lobes. Especially in ISAR imaging, large gapped data will lead poor quality of the imagery with the above method. The problem of interest in this paper is how to obtain the perfect image with incomplete aperture.

Now let’s take a look at the transmitted signal, a FMCW signal as:

$$s(\tau) = A \cdot rect\left(\frac{\tau}{T_p}\right) \cdot e^{j2\pi f_c \tau + j\pi \gamma \tau^2}$$  

(5)

where $A$ is the amplitude, $f_c$ is the carrier frequency, $\gamma$ is the chirp rate, $T_p$ denotes time width of the chirp pulse and $rect()$ stands for the unit rectangular function. After illuminating the target scene, the echo signal from the object should be:

$$r(\tau, t_m) = A(t_m) \cdot rect\left(\frac{\tau}{T_p}\right) \cdot rect\left(\frac{t_m}{T_a}\right) \cdot e^{j2\pi [f_c(\tau - \frac{2R(t_m)}{c}) + \frac{\gamma}{2}(\tau - \frac{2R(t_m)}{c})^2]}$$  

(6)

where $c$ is the speed of light, $R$ is the instantaneous distance between radar and target. The second figure in Figure 4 shows the sketch of the transmitting signal and echoed signals. After range compression, translation compensation and phase adjudgment processes, the signal in a some range cell is:

$$s(t_m) = \sum_{p=1}^{P} A_p \cdot rect\left(\frac{t_m}{T_p}\right) \cdot e^{-j2\pi f_p t_m}$$  

(7)

where $P$ is the number of scatters in range cell and $T_p$ is coherent processing interval, $f_p$ is Doppler of the scatter.
4 Data Recovery by Compressed Sensing

In order to apply the theory of CS to radar imaging, we should firstly construct a measurement matrix $\Phi$. It is proved that random matrix performs well, such as Gaussian or Bernoulli matrix. Next step is how to construct the basis matrix $\Psi$. We should note that there is no need for CS to sample the signal with constant period, so we construct the basis matrix as:

$$
\Psi = \begin{bmatrix}
\exp(2\pi j f_1 t_0) & \exp(2\pi j f_2 t_0) & \cdots & \exp(2\pi j f_N t_0) \\
\exp(2\pi j f_1 t_1) & \exp(2\pi j f_2 t_1) & \cdots & \exp(2\pi j f_N t_1) \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
\exp(2\pi j f_1 t_M) & \exp(2\pi j f_2 t_M) & \cdots & \exp(2\pi j f_N t_M)
\end{bmatrix}
$$  \hfill (8)

Scatters are separated for different Doppler frequency in range cell in ISAR system, and the key point is to how to find the frequency and amplitude of the object for sparsity signal. Donoho et al. [Don06] introduced a solution for recovery of compressed signal by convex optimization as:

$$\min \|s\|_1 \text{ s.t. } \|y - \Theta s\|_2 \leq \epsilon$$  \hfill (9)

where $\epsilon$ is the error.

Tropp et al. [TG07] proposed Basis Pursuit (BP) and Orthogonal Matching Pursuit (OMP) algorithms are commonly used. In our system, we can use it as:

$$\min \|f_d\|_1 \text{ s.t. } \|y - \Theta f_d\|_2 \leq \epsilon$$  \hfill (10)

Furthermore, as shown in paper [LDP07], we can further improve the convex optimization function in this way:

$$\min \|f_d\|_1 + \alpha TV(s) \text{ s.t. } \|y - \Theta f_d\|_2 \leq \epsilon$$  \hfill (11)

where $\alpha$ is weighted coefficient and trades sparsity with finite-differences sparsity, and $TV$ means Total Variance. The $l_1$ norm in the objective is a crucial feature of the whole approach. Minimizing the $l_1$ norm of an objective often results in a sparse solution. On the other hand, minimizing the $l - 2$ norm is commonly used for regularization because of its simplicity, which does not result in a sparse solution and hence is not suitable for use as objective function. Intuitively, the $l - 2$ norm penalizes large coefficients heavily, therefore solutions tend to have many smaller coefficients – hence not be sparse. In the $l_1$ norm, many small coefficients tend to carry a larger penalty than a few large coefficients, therefore small coefficients are suppressed and solutions are often sparse.

With signal processing and convex optimization as above, we can recover all the signal data. Next we can recover the image of aperture imaging system from the recovered data. The whole imaging procedure can be seen from Figure 4. The procedure for digital spotlighting can be summarized as follows: First we need to calculate the reference SAR signal $s_0(\omega, u)$. And then polar format process with the digital spotlight filter. After that we obtain the digital-spotlighted slow-time compressed SAR signal in the $(\omega, k_u)$ domain by Fourier Transform. After optional zero-padding in the $k_u$ domain, the slow-time compressed signal can be formed by an inverse Fourier Transform. Finally the digital spotlight SAR signal is constructed from matched filtering with the reference SAR signal. In next section we will discuss how to speed up this process by applying sparse FFT algorithm.
5 Sparse FFT Image Reconstruction

We use the technique from [HIKP12a] to speed up the Fourier Transform of Image Reconstruction.

A general algorithm for computing the exact DFT must take time at least proportional to its output size, i.e., $\Omega(n)$. However, most of the Fourier coefficients of a image/audio signal are small or equal to zero, i.e., the output of the DFT is (approximately) sparse.

If a signal has a small number $k$ of non-zero Fourier coefficients, the output of the Fourier transform can be represented succinctly using only $k$ coefficients. Hence, for such signals, one may hope for a DFT algorithm whose running time is sublinear in the signal size, $n$. The goal of sparse Fourier Transform is to find an approximate vector $\hat{x}'$ that satisfies the $l_2/l_2$ guarantee with sublinear time.

At a high level, sparse Fourier algorithms work by binning the Fourier coefficients into a small number of bins. Since the signal is sparse in the frequency domain, each bin is likely to have only one large coefficient, which can then be located (to find its position) and estimated (to find its value). We use some good property filter to manipulate signal in time domain which can be seen as a binning in frequency domain. In order to get a sublinear time algorithm, our filter in time domain should have sublinear support size while its approximate corresponding frequency domain filter should be like a box filter in order to separate the frequency domain into small bins.

A good flat filtering window is $sinc \times gaussian$ in time domain or $box \ast gaussian$ in frequency domain. With these filter we can use sublinear time algorithm to manipulate the time domain same as binning the frequency domain. After that we can take sample from the frequency domain to recover the approximating location and value based on sampling from each bin.

Another thing during binning we need to care about is randomness. Since the support of
frequency domain vector \( \hat{x} \) may lie in a small area and make our binning difficult to separate, we need to first permute the frequency domain. Fortunately, this random permutation can be done with tools called spectrum permutation which have 2-way independent with parameter \( a, b \).

If we only discuss about exact k-sparse recovery, the operation above only guarantee recovering \( k/2 \) of the vector in frequency domain if the frequency domain is a k-sparse vector. Hence it takes us \( O(k \log n) \) to approximately recover, where our result \( \hat{z} \) satisfies \( \hat{x} - \hat{z} \) is k/2-sparse. After that we can continue our sparse recovery on \( \hat{x} - \hat{z} \), and since it is k/2 sparse we can do with it with \( O(k/2 \log n) \) time. This would give us a vector such that its difference with \( \hat{x} \) is k/4 sparse. Repeated the procedure until we reach constant sparsity then we can say we exact recover for \( k - \text{sparse} \) signal \( x \). Notice that the total time is still \( O(k \log n) \) because \( k + k/2 + k/4 + ... < 2k \).

We haven’t finished our implementation of sparse FFT yet. The most difficult part is the construction of the flat filtering windows. We need to construct them efficiently with a good parameter that can compute fast in time domain and exactly fit with the binning number \( B \) we need. Another limitation is that even with sparse FFT, only when the ratio of size of data \( n \) and the sparsity \( k \) \( n/k \) is larger than 2000. However in real application of SAR, we can get such large size signal.

6 System Implementation

In this project, we are aimed to develop an acoustic imaging system based on the smartphone and only use built-in speakers and microphones. The advantage of this setup is clear: the user does not need to bring any extra equipment for acoustic imaging, which makes it truly ubiquitous. Based on initial exploration, we applied PFA algorithm to acoustic aperture imaging. In this case, we move the target smartphone along a certain trajectory and the microphone on smartphone is used to record the acoustic signals. Also, we need to apply RMA to remove the near field effect (i.e., the edges of the object will become bended when the receiver and the targeted target is close to each other).

However, our system still introduces several challenges for acoustic imaging. First, it is hard to move a smartphone by hand to precisely follow a pre-defined trajectory. The trajectory error will cause the phase of the received signals to deviate from the desired values. Due to the short wavelength \( \lambda \) of the acoustic signal (around 2 cm), even a small trajectory error \( \delta_d \) will lead to a large phase error \( 2\pi \delta_d / \lambda \). To generate a clear image of the object, these phase errors need to be removed. Second, the speakers on the smartphone are for general purpose usage and cannot generate ultrasound. The highest frequency supported is usually about 20 KHz - 22 KHz. Moreover, to avoid the generated sound signals audible by the user, its frequency should be higher than 16 KHz. As a result, the all available frequency is from 16 KHz - 22 KHz, i.e., the usable bandwidth of the acoustic signal is about 6 KHz. The proposed imaging approach should provide reasonable resolution with the limited bandwidth. Third, the acoustic signals produced by the speakers on the smartphone propagate over all directions, instead of concentrating onto the targeted objects. As a result, the received signals contain not only the reflections from the target objects, but also those from irrelevant ones. The existence of these irrelevant reflections will interfere with the desired received signals, and blur or even overwhelm the generated images of the target object. Effective mechanisms need to be developed to mitigate such interference.

In order to fix the problems and challenges described above, we will do some simulations to check whether our proposed approaches can recover the sparse aperture signal and improve the imaging resolution.
7 Simulation and Evaluation

In our simulation setup, we define the audio propagation speed \( c = 343.2 \text{ m/s} \); baseband bandwidth \( f_0 = 4k \); carrier frequency \( f_c = 19k \); range distance to center of target area \( X_c = 1m \); target area radius in range \( X_0 = 0.5m \); cross-range distance to center of target area \( Y_c = 0 \); target area radius in cross-range \( Y_0 = 0.5m \); synthetic aperture is \( 2 \times L = 0.4m \); chirp pulse duration is \( 10ms \). There are initialed 9 target points in the image.

Let’s first see the signal processing results of each step with sparse aperture data.

![Figure 5: Measured Signal.](image1)

![Figure 6: Fast-time MF.](image2)

![Figure 7: Aliased Sig.](image3)

![Figure 8: Compressed Sig.](image4)

![Figure 9: Polar Format Reconstruction in Spotlight.](image5)

![Figure 10: DS Upsampling.](image6)

![Figure 11: Polar Format.](image7)

![Figure 12: DS Upsampling.](image8)

![Figure 13: Compressed Spect.](image9)

We can see that there are nearly half of the aperture data are missed from Figure 5. Figure 6 shows the SAR signal after fast-time Matched Filtering, where there are several blur lines. Figure 7 shows the aliased signal spectrum. Figure 8 -15 shows the signal spectrum after compression, polar format and up-sampling processes. Finally Figure 16 - 13 illustrate the spotlight reconstruction spectrums.
Finally let’s see the sparse aperture recovery result after compressed sensing. We compare four different recovery strategies illustrated in section 4. We can see that our compressed sensing approaches can make big differences on imaging recovery resolution.

![Figure 14: Zero-padding.](image1)

![Figure 15: Donoho CS Opt.](image2)

![Figure 16: BP & OMP.](image3)

![Figure 17: $l_1 + l_2$ norm Opt.](image4)

### 8 Conclusion and Future Work

In our project, we propose a novel mobile framework for ISAR imaging system. We next recover signal from sparse aperture in azimuth bin and propose a new method for sparse aperture imaging on mobile device. In addition, we intend to speed up Fourier Transform image reconstruction process and use only a subset of the input data required by the FFT, but haven’t finish this part. Simulation results show the advantages of our CS approach that we can get finer imaging resolution with fewer valid aperture data.

Many future work can be done to further improve the performance of our system. First is to finish the remaining algorithms. In addition, apply Sparse FFT to align two cancellation images from background signal and target signal to make sure that the corresponding pixels represent the same position in the real world. Furthermore, we can try to apply Phase Gradient Auto-focus (PGA) algorithm [WEGJ94] to iteratively estimate and compensate the phase error from the trajectory error.
9 Acknowledgment

This is the final report of course project for 2016 Fall UT-Austin graduate course CS 395T Sublinear Algorithm. Thanks a lot for the help from Prof. Eric Pice and TA Zhao Song.
References


