Design and Performance Evaluation of an Indoor Ranging System Using Audio Chirp and MUSIC Algorithms

HAMID FARROKHI
Department of Engineering, Birjand
University, P.O. Box 79, Birjand
IRAN
hfarrokhi@birjand.ac.ir http://www.birjand.ac.ir

Abstract: - In most existing indoor acoustic ranging systems, a range is estimated based on measuring the TOA of an ultrasonic wave. Due to high directivity of ultrasonic transducers and high attenuation of ultrasonic waves with distance in air, these systems cannot provide a range in large indoor environments. To achieve longer ranges, a TOA-based acoustic ranging system that uses a low frequency audible signal is proposed. The audible sound is a spread spectrum LFM signal, which is generated with a standard speaker and received with a standard microphone. The indoor acoustic environment is modeled with a multipath propagation channel whose characteristics are different from those of outdoor environments. It is shown that the classical ranging techniques used in traditional systems, such as the GPS, Sonar, Radar, etc., may not provide the optimum performance in indoor environments. The complexity and site-specific nature of indoor acoustic propagation, when using traditional correlation techniques, often lead to the creation of undetectable direct-line-of-sight situations. It is shown in this paper that by using MUSIC super-resolution techniques, namely the root-MUSIC, which is applied to the estimated frequency response of the multipath indoor propagation channels, the substantial TOA errors can be reduced significantly and the range accuracy improved by a factor of at least 3.

Key-Words: - indoor localization, chirp, root-MUSIC, LFM.

1 Introduction and Problem Definition

The traditional RF positioning systems, such as the GPS, are not designed for indoor environments where the DLOS between the receiver and the satellite does not exist. On the other hand, there is a great interest in location-based applications which require accurate position estimation for these environments. So, there is a need for another positioning technique and system such as the acoustic ranging system [1] shown in figure 1. In such a system an audible chirp pulse $s(t)$, which is a spread spectrum (SS) linear frequency modulated (LFM) sound signal emanates from a speaker. The signal is monitored and collected by a microphone placed in a spatially separated receiver so that it estimates the propagation delay of the received signal. The time of arrival (TOA) is apparently the most widely used metric in acoustic ranging systems [2] and can be readily derived, during each measurement, from the impulse response, $h(t)$ of indoor propagation channel. In literature, models presented for an indoor acoustic channel are usually very complicated [1] and the impulse response is not fixed, but is a site-specific and time-variant function of channel parameters such as temperature, speed of sound, air currents, relative humidity, etc. It should be noted that for small measurement time snapshots these parameters can be considered as constant and the model is assumed to be linear and time invariant. In this case, $h(t)$ can be mathematically written as a summation of weighted delta functions [1], that is

$$h(t) = \sum_{k=0}^{L_{p-1}} \alpha_k \delta(t - \tau_k) \quad (1)$$

where L_p is the number of multipath components, α_k and τ_k are complex amplitude and propagation delay of the kth path, respectively. Figure 2 depicts such an impulse response, in which τ_0 is the delay

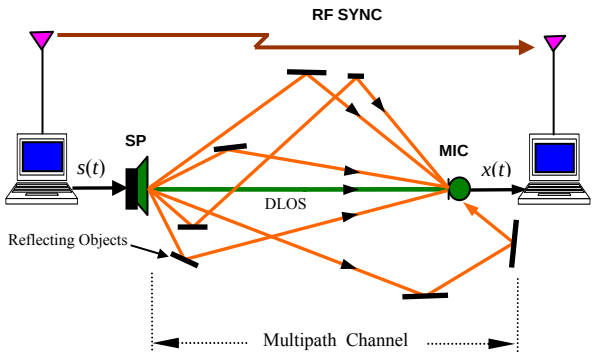

Figure 1. An RF-Synchronized Indoor Acoustic Ranging System

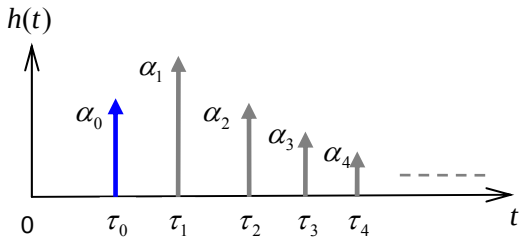

Figure 2. Impulse response of indoor channel

corresponding to the DLOS path and needs to be estimated as the TOA. As a result, the received signal is a summation of the DLOS and superimposed multipath signals reflected from surrounding objects all arriving the microphone with different delays and amplitudes, and the additive noise. It is modeled as

$$x(t) = s(t) * h(t) = \sum_{0}^{L_p - 1} \alpha_k s(t - \tau_k) + n(t) \qquad (2)$$

where $s(t)$ is the transmit signal, $n(t)$ is assumed to be additive white Gaussian noise (AWGN). A radio signal, indicated as RF SYNC in Figure 1, can be used for the purpose of synchronization between the receiver and the transmitter. TOA is determined at the receiver by taking the time difference between the arrivals of the RF SYNC and the audible chirp, $x(t)$. This is accomplished by using a matched filter (MF) that maximizes the peak signal-to-noise ratio in the presence of additive white noise. In the absence of Doppler distortion, the MF operation is reduced to the cross-correlation function (CCF) of the received signal with a copy of the transmitted signal[1]. The output of the matched filter, $r_{xs}(\tau)$, is a weighted sum of multiple shifts of the auto-correlation function (ACF) of $s(t)$, and is given by [2]

$$r_{xs}(\tau) = \sum_{k=0}^{L_p - 1} \alpha_k r_{ss}(\tau - \tau_k) + u(\tau) \qquad (3)$$

in which, $u(\tau)$ is the output additive noise term. The time delay $\tau_0 (k = 0)$ corresponding to the first peak of $r_{xs}(\tau)$ is taken as an initial estimate for TOA and depends on the shape of the ACF. The function $|x_{xs}(t)|$ is called the delay profile. Typical delay profiles are shown in Figure 3. It is observed that when the difference between the adjacent path delays is greater than the width of the ACF, W, then clearly separated correlation peaks appear in the delay profile (Figure 3(a, b)) so that the TOA can be determined by finding the delay value corresponding to the first peak. If the strength of the first peak is below the receiver sensitivity or the difference between the amplitudes of the first and the next peaks exceeds the receiver dynamic range, the DLOS path is not detectable. In this case, dramatically large TOA estimation errors may occur

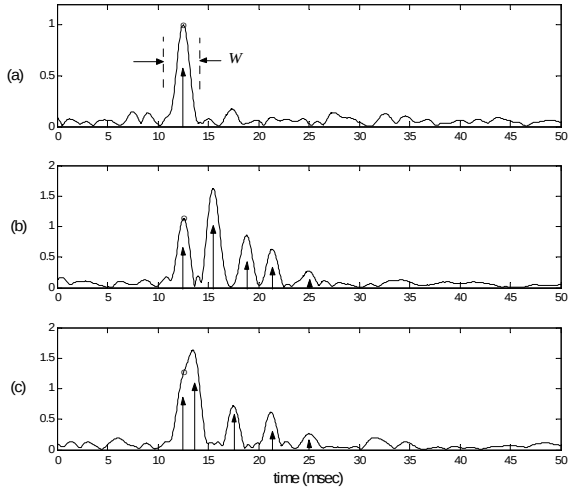

Figure 3. Example of multipath delay profiles.

Another situation, in which substantial range error may occur is when the time difference between the adjacent delays is smaller than W and the two delayed ACFs overlap. Due to the construction and deconstruction effects between the two overlapped ACFs, as shown in Figure 3(c), large estimation errors occur if the TOA is estimated by detecting the first peak.

Therefore, the problem with the traditional TOA-based systems in indoor multipath environments is to accurately estimate the propagation delay of the signal arriving from the direct-line-of-sight (DLOS). In multipath channels the following general principles can be used to improve the performance of the correlation method in TOA estimation [2]:

1. Increasing the receiver's dynamic range and sensitivity especially for situations, in which the DLOS is detectible.
2. Increasing the signal bandwidth to reduce the overlap between autocorrelation functions.
3. Employing high-resolution techniques to resolve for the TOA.

The last two principles are implemented for scenarios where the DLOS exists but is not detectible. Increasing the signal bandwidth, B, reduces the spread of the correlation peak and helps to resolve the DLOS path.

In the proposed acoustic ranging system, although a large amount of bandwidth is available in the

[1]In the presence of Doppler distortion in situations where the receiver acts as a mobile terminal, the MF output is the ambiguity function.

audible frequency range, but in some indoor SS acoustic ranging applications, one might not use the whole available bandwidth due to the following criteria:

- Avoiding the use of large frequency bandwidth to cope with attenuation of sound at high frequencies [3], especially for long distance ranging.
- Increasing the covered area due to transducer's directivity. Since the transducer's directivity increases with frequency, the area covered by the ranging system decreases.
- Dealing with occlusions or obstructed-line-of-sight (OLOS) scenarios [10], in which the high frequency components of the DLOS acoustic wave cannot reach the receiver and are totally reflected by the occlusions.

In addition, achieving higher ranging accuracy using the same frequency bandwidth is always desirable.

Inspired by high resolution TOA estimation techniques for RF-based systems in [2, 4, 11], a frequency-domain super-resolution technique can be used to determine the TOA with high resolution from the estimated frequency channel response.

Estimation of the TOA falls into the category of signal parameter estimation. The single-path model for propagation channel and traditional TOA estimation techniques, namely the correlation techniques used in the outdoor positioning systems, such as GPS, Radar, and Sonar, might not provide an accurate estimate of a range. Therefore, significant amount of works has been devoted in this paper to the implementation of a higher resolution technique such as MUSIC algorithms. These algorithms are applied to the frequency-domain data, which are samples of the estimated frequency response of the indoor multipath propagation channel, in order to more accurately estimate the TOA.

2 High-resolution TOA Estimation-MUSIC Algorithms

As it was mentioned in the previous section, the performance of TOA estimation can be improved by using high-resolution techniques [4, 5]. Among different high-resolution methods, the Multiple Signal Classification (MUSIC) algorithms have been used in this paper. These algorithms belong to an eigen-decomposition-based category of super-resolution techniques for estimating the frequencies of superimposed sinusoid signals in noise.

Using Eq. (1), the frequency-domain channel response in the absence of noise is calculated as the following superimposed L_p complex sinusoids given by

$$H(f) = \sum_{k=0}^{L_p-1} \alpha_k e^{-j2\pi f \tau_k} \tag{4}$$

In Eq. (4), $H(f)$ is the Fourier transform of $h(t)$ and $\alpha_k = |\alpha_k| e^{j\theta_k}$, in which θ_k is the phase of the complex attenuation of the kth multipath arrival. The parameters α_k and τ_k are generally random time-variant functions because of the motion of the receiver or other acoustic reflecting objects or temporal variations in the physical characteristics of an environment such as air currents or temperature gradients.

The MUSIC algorithms require, as their input data, an estimate of the propagation transfer function $H(f)$ in the form of discrete samples at L equally spaced frequencies. If noise is AWGN and the estimated transfer function in noise is denoted by $\hat{H}(f)$, then the sampled discrete transfer function of the channel is given by [2]

$$\begin{aligned} y(l) &\triangleq \hat{H}(f_l) = H(f_l) + w(l) \\ &= \sum_{k=0}^{L_p-1} \alpha_k e^{-j2\pi f_l \tau_k} + w(l) \\ &= \sum_{k=0}^{L_p-1} \alpha_k e^{-j2\pi (f_0 + l\Delta f)\tau_k} + w(l) \end{aligned} \tag{5}$$

where $l = 0,1,\ldots,L-1$, $w(l)$ denotes the additive white measurement noise with zero mean and variance σ_w^2, and f_0 and Δf are the start frequency and the frequency-domain sampling interval, respectively. Eq. (5) denotes a summation of complex sinusoids. The vector form of Eq. (5) is

$$\mathbf{y} = \mathbf{H} + \mathbf{w} = \mathbf{V}\mathbf{a} + \mathbf{w} \tag{6}$$

where

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} y(0) & y(1) & \cdots & y(L-1) \end{bmatrix}^T \\ \mathbf{H} &= \begin{bmatrix} H(f_0) & H(f_1) & \cdots & H(f_{L-1}) \end{bmatrix}^T \\ \mathbf{w} &= \begin{bmatrix} w(0) & w(1) & \cdots & w(L-1) \end{bmatrix}^T \\ \mathbf{V} &= \begin{bmatrix} \mathbf{v}(\tau_0) & \mathbf{v}(\tau_1) & \cdots & \mathbf{v}(\tau_{L_p-1}) \end{bmatrix} \\ \mathbf{a} &= \begin{bmatrix} \alpha_0' & \alpha_1' & \cdots & \alpha_{L_p-1}' \end{bmatrix}^T, \end{aligned} \tag{7}$$

in which

$$\mathbf{v}(\tau_k) = \begin{bmatrix} 1 & e^{-j\,2\pi\Delta f\,\tau_k} & \cdots & e^{-j\,2\pi(L-1)\Delta f\,\tau_k} \end{bmatrix}^T \quad (8)$$

and

$$\alpha_k^{'} = \alpha_k e^{-j\,2\pi f_0 \tau_k} \quad (9)$$

and the superscript T denotes the matrix transpose operation.

The main objective of these algorithms is to estimate the frequencies of the complex sinusoids in noise by exploiting an eigenvalue decomposition of the correlation matrix of the discrete channel model in Eq. (6), which is given by

$$\mathbf{R}_{\mathbf{yy}} = E\{\mathbf{y}\mathbf{y}^H\} = \mathbf{V}\mathbf{A}\mathbf{V}^H + \sigma_w^2\mathbf{I} \quad (10)$$

where

$$\mathbf{A} = E\{\mathbf{a}\mathbf{a}^H\} \quad (11)$$

and $\mathbf{I}$ is the identity matrix. The superscript H denotes the conjugate transpose operation, i.e., the Hermitian of a matrix. Since the propagation delays τ_k are all different, then the column vectors of the signal matrix $\mathbf{V}$ are linearly independent, i.e., the matrix $\mathbf{V}$ is always full rank (L_p). If we assume that α_k is random, the $L_p \times L_p$ correlation matrix $\mathbf{A}$ is non-singular. It follows that the rank of $\mathbf{V}\mathbf{A}\mathbf{V}^H$ is L_p (assuming $L > L_p$). Let the eigenvalues of $\mathbf{V}\mathbf{A}\mathbf{V}^H$ be $\{v_0 \ge v_1 \ge \cdots \ge v_{L_p-1}\}$ and those of $\mathbf{R}_{\mathbf{yy}}$ be $\{\lambda_0 \ge \lambda_1 \ge \cdots \ge \lambda_{L-1}\}$. According to Eq. (4.8) we have [2]

$$\lambda_i = \begin{cases} v_i + \sigma_w^2, & i = 0,\ldots,L_p - 1 \\ \sigma_w^2, & i = L_p,\ldots,L-1 \end{cases} \quad (12)$$

which means that the eigenvalues (and eigenvectors) may be partitioned into two sets, those that correspond to the signals and those that correspond to noise. Let $\mathbf{Q} = \{\mathbf{q}_0, \mathbf{q}_1, \cdots, \mathbf{q}_{L-1}\}$ be the eigenvectors of $\mathbf{R}_{\mathbf{yy}}$. All the $L - L_p$ eigenvectors associated with the smallest eigenvalues of $\mathbf{R}_{\mathbf{yy}}$ satisfy the relations

$$\mathbf{R}_{\mathbf{yy}}\mathbf{q}_i = \sigma_w^2\mathbf{q}_i \ , \ i = L_p, \cdots, L-1 \quad (13)$$

or, equivalently

$$(\mathbf{R}_{\mathbf{yy}} - \sigma_w^2\mathbf{I})\mathbf{q}_i = \mathbf{0} \ , \ i = L_p, \cdots, L-1 \quad (14a)$$

Using Eq. (10), Eq. (14) may be rewritten as

$$\mathbf{V}\mathbf{A}\mathbf{V}^H\mathbf{q}_i = \mathbf{0} \ , \ i = L_p, \cdots L-1 \quad (14b)$$

Since the matrix $\mathbf{V}$ is full rank L_p and matrix $\mathbf{A}$ is diagonal with all nonzero entries, it follows from Eq. (14b) that

$$\mathbf{V}^H\mathbf{q}_i = \mathbf{0} \ , \quad i = L_p, \cdots L-1 \quad (15)$$

which means that the columns of $\mathbf{V}$ are related to the eigenvectors corresponding to the $L - L_p$ eigenvalues of $\mathbf{R}_{\mathbf{yy}}$ by [12]

$$\mathbf{v}^H(\tau_k)\mathbf{q}_i = 0 \ , \quad \begin{matrix} i = L_p, \cdots, L-1 \\ k = 0, \cdots, L_p - 1 \end{matrix} \quad (16)$$

and the continuous variable time scanning vector, $\mathbf{v}(\tau)$, is defined as

$$\mathbf{v}(\tau) = \begin{bmatrix} 1 & e^{-j\,2\pi\Delta f\,\tau} & \cdots & e^{-j\,2\pi(L-1)\Delta f\,\tau} \end{bmatrix}^T \quad (17)$$

In order to avoid aliasing in the time-domain, similar to the time-domain Nyquist sampling theorem, the value of Δf is determined so as to satisfy the condition

$$\Delta f \le \frac{1}{2\tau_{\max}} \quad (18)$$

where, $\tau_{\max} = \tau_{L_p-1}$. is the maximum delay of the measured multipath in the acoustic propagation channel.

The MUSIC algorithms including spectral MUSIC and root-MUSIC exploit Eq. (16) for the purpose of determining the propagation delays, τ_k, $0 \le k \le L_p - 1$. In the spectral method, the delays, τ_k, $0 \le k \le L_p - 1$, can be determined by finding the time delays at which the pseudo-spectrum $S_{MUSIC}(\tau)$ given by [2]

$$S_{MUSIC}(\tau) = \frac{1}{\sum_{k=L_p}^{L-1} \left| \mathbf{q}_k^H \mathbf{v}(\tau) \right|^2} \quad (19)$$

achieves its maximum values. The estimate of TOA is then determined by detecting the first peak of the pseudospectrum along the delay axis. The time scanning vector defined in Eq. (17) should scan the complete time interval, $0 \le \tau \le 1/2\Delta f$. Outside this interval the pseudospectrum repeats itself. Figure 2 shows a sample result of the time-domain MUSIC pseudospectrum obtained from the indoor acoustic ranging system measurement.

The need for time scanning can be avoided by using a root-finding approach [6, 8]. In this approach, the complex exponential, $e^{j\,2\pi\Delta ft}$, is replaced by the

complex variable, z, in Eq. (17). If $D(z)$ denotes the resulting denominator polynomial, then

$$S_{MUSIC}(z) = \frac{1}{D(z)} \quad (20)$$

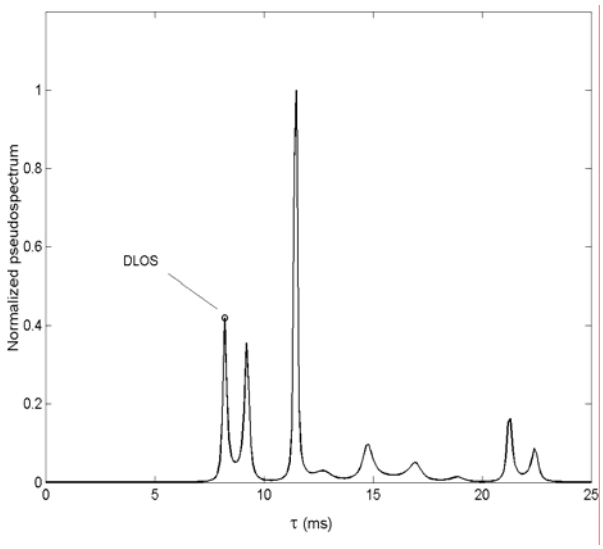

Figure 4. A time domain MUSIC pseudospectrum.

The roots of the polynomial in Eq. (20) are used to estimate the propagation delays, τ_k, $0 \le k \le L_p - 1$. The L_p roots, $z_k = |z_k| e^{j\beta_k}$, $0 \le k \le L_p - 1$, of $D(z)$, which lie on the unit circle or closest to it, correspond to signals and the phases of these roots (in radians) are related to the propagation time delays by [8]

$$\tau_k = \frac{\beta_k}{2\pi\Delta f} \quad (21)$$

Details of the root-MUSIC can be found in [8]. Figure 4 shows roots of $D(z)$ as a result of the root-MUSIC algorithms for a 3-path channel with L = 14. In this case, $D(z)$ has a total of 28 roots. The root corresponding to the DLOS component, which is shown as the first arrow after the 0-degree axis, and other multipath signal components must lie on the unit circle or closest to it. It is worth noting that each root of $D(z)$ has its inverse located outside the unit circle.

In [8], there are simulation results that consider the effect of an error Δz_i in the signal zeros, $z_i = |z_i| e^{j\beta_i}$, that show that the spectral MUSIC and root-MUSIC have the same asymptotic mean squared errors, but root-MUSIC is preferable. According to the simulations, if an error Δz_i is

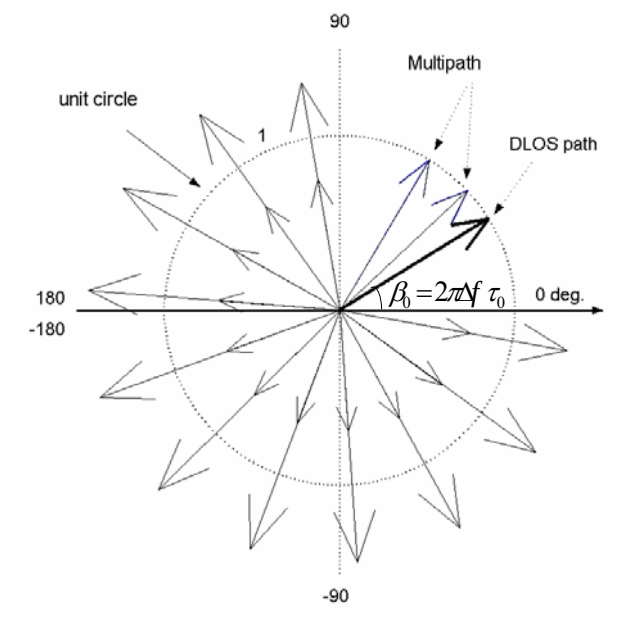

Figure 5. Root-MUSIC diagram for a 3-path channel ($L_p = 3$) with L = 14.

radial, then there is no error in the estimate of the phase $\beta_i = 2\pi\Delta f\ \tau_i$ and consequently, in the time delay τ_i. However, such radial errors do affect the spectral MUSIC by making the peaks in $S_{MUSIC}(\tau)$ less defined. This is particularly critical for closely spaced roots as they may result in only one peak causing an apparent loss in resolution. So, spectral methods always have less resolution compared to root forms. To understand the degree to which this is true, one needs to study the effects of both Δz_i and $\Delta\beta_i$. An important outcome of the analysis and simulations in [8] is the fact that an error Δz_i in a signal zero z_i has a largely radial component. This provides an explanation as why root-MUSIC is superior to the spectral MUSIC algorithm. The TOA estimation method in this work is based on the root-MUSIC algorithm.

3 Issues of Practical Implementation

In the last section, which led to the spectral MUSIC and root-MUSIC TOA estimation, the estimation of the channel frequency response and the correlation matrix were required to calculate the propagation time delays. In practice, since the channel frequency response $H(f)$ and as a result, the correlation matrix is not consistent for consequent measurements, the ranging system should have an estimate of the channel frequency response and the correlation matrix for each data measurement snapshot. Figure 2 illustrates a functional block diagram of a super-resolution system which operates

based on MUSIC algorithms. The input data vector, i.e., the estimate of channel frequency response given in is first used to estimate the correlation matrix $\mathbf{R}_{yy}$. Then the eigenvalues and the corresponding eigenvectors of the correlation matrix are computed. The parameter L_p is determined through the analysis of the eigenvalues and eigenvectors of the correlation matrix. Finally, either the pseudospectrum $S_{MUSIC}(\tau)$ or the signal zeros z_i are obtained.

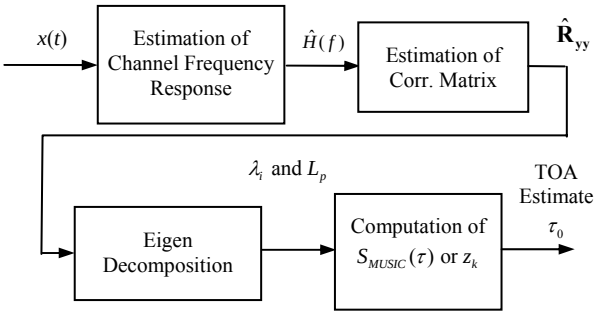

Figure 6. Functional block diagram of MUSIC algorithms

3.1 Δf and Dimensions of $\mathbf{R}_{yy}$

It was mentioned in section 2 that in order to avoid aliasing in the time-domain, similar to the time-domain Nyquist sampling theorem, the frequency-domain sampling interval, Δf , is determined so as to satisfy the Eq. (18) [13]. For example, suppose that the maximum length of the multipath acoustic signal propagation path is 50 m. If the speed of sound, $c = 345$ m/s, then the corresponding maximum delay, $\tau_{\max} = 145$ ms. Using (18) for the frequency-domain sampling interval, we obtain $\Delta f = 3.45$ Hz.

In TOA estimation using acoustic waves, the low propagation speed of sound (compared to RF) results in large values of time delays. According to Eq. (18), a large value of the time delay leads to a small value for Δf . On the other hand, the length N of the measurement data equals the ratio of the signal bandwidth and the frequency sampling interval, Δf . In the above example, if the size of the correlation matrix, L, is chosen such that $L = (2/3)N$, and the signal bandwidth, B, is set to 2 kHz, then N = 580 and L = 387. An immediate consequence is that the super-resolution techniques for TOA estimation demand high calculation time and complexity of practical implementation, as we might have to deal with correlation matrices of large dimensions. In fact, it can almost be impractical, especially for long range measurements, in which $\tau_{\max}$ has a large magnitude. One solution to this problem is to first obtain a coarse estimate of TOA and then apply the super-resolution algorithm only to the informative part of the received signal, which can be isolated from the non-informative part. This can be accomplished by including two additional steps of correlation and windowing. The correlation step is used to find a raw estimate of TOA, and the windowing step is required for isolating the informative part of the correlation waveform, respectively. By using this method, the time delay corresponding to the non-informative part of the received signal (t_d in Figure 11) is not considered in the super-resolution process and as a result, the size of the correlation matrix will be reasonable. The details regarding this method will be presented in Sec. 3.6.

3.2 Signal De-correlation

Throughout the MUSIC algorithms, it is assumed that the signals to be estimated are uncorrelated, which means that the parameters, α_k, $0 \le k \le L_p - 1$, in Eq. (4) are random[1] [2]. With such assumptions the signal components in the vector $\mathbf{a}$ in Eq. (7) are non-coherent, *i.e.*, uncorrelated and the $L_p \times L_p$ correlation matrix, $\mathbf{A}$, defined in Eq. (11), is nonsingular. Also the correlation matrix of the data, $\mathbf{R}_{yy}$, is Hermitian, *i.e.*, conjugate symmetric, and Toeplitz, *i.e.*, having equal elements along all diagonals. In practice, however, the data measurement time snapshots are usually small as compared to the rate of variations in the parameters. So, the parameters, α_k, $0 \le k \le L_p - 1$, are non-random during one data measurement snapshot. On the other hand, if these parameters do not change with frequency over the bandwidth of the acoustic signal, they can be considered as constant ,but unknown, which means that in multipath environments, the signal components are coherent (correlated) and the matrix, $\mathbf{A}$, defined in Eq. (11) is singular and rank deficient and assumption in Eq. (15) no longer holds. The coherent signals in a multipath environment can have detrimental effects on the performance of MUSIC [17]. Therefore, in order to use the MUSIC techniques for indoor TOA estimation, we need to 'de-correlate' the received signals. There are two techniques for destroying the signal coherence [4]. One is spatial smoothing

[1] It can be assumed that $|\alpha_k|$, $0 \le k \le L_p - 1$, are constant and θ_k, $0 \le k \le L_p - 1$, are uniform random variables in $[0,\ 2\pi]$.

preprocessing (SSP). The other is modified spatial smoothing preprocessing (MSSP). These two techniques are discussed in sequel.

It is assumed that only one snapshot of the measurement data of length N is available, which is the case in the proposed system. In the SSP technique, the data sequence is divided into M consecutive and overlapping segments, or sub-arrays, of length L and then the correlation matrix is estimated by sliding the sub-array across the full-array of data points in the forward direction. Then, the correlation matrix is estimated by

$$\hat{\mathbf{R}}_{\mathbf{yy}} = \frac{1}{M}\sum_{k=0}^{M-1}\mathbf{y}(k)\mathbf{y}(k)^H \qquad (22)$$

where

$$\mathbf{y}(k) = \left[y(k) \quad \ldots \quad y(k+L-1)\right]^T \qquad (23)$$

and

$$M = N - L + 1 \qquad (24)$$

This method is illustrated in Figure 4, in which the M overlapped segments are used for calculating the correlation matrix. The correlation matrix obtained by Eq. (22) is called the 'forward correlation matrix' (FCM) [4]. For actual measurement data, the FCM defined in Eq. (22) may not be Toeplitz.

In the MSSP technique, de-correlation is introduced by sliding the sub-array in both the forward and backward directions. By the use of this technique, as we will see later, more improvement in the estimation of the correlation matrix can be achieved, when compared to the SSP technique. The resulting correlation matrix is called the 'forward-backward correlation matrix' (FBCM) and is given by [4]

$$\hat{\mathbf{R}}_{\mathbf{yy}}^{(FB)} = \frac{1}{2M}\sum_{k=1}^{M}\left(\hat{\mathbf{R}}_{\mathbf{yy}} + \mathbf{J}\hat{\mathbf{R}}_{\mathbf{yy}}^{*}\mathbf{J}\right) \qquad (25)$$

where $\hat{\mathbf{R}}_{\mathbf{yy}}$ is defined in Eq. (22) and $\mathbf{J}$ is the $L \times L$ exchange matrix, defined as

$$\mathbf{J} = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & \iddots & 1 & 0 \\ 0 & \iddots & \iddots & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \qquad (26)$$

Although the resolution is reduced due to the smoothing, yet the MUSIC algorithms using the preprocessing techniques realizes much better performance than conventional techniques [4]. The de-correlation effects in FCM and FBCM matrices have been analyzed in [4, 13] by calculating the

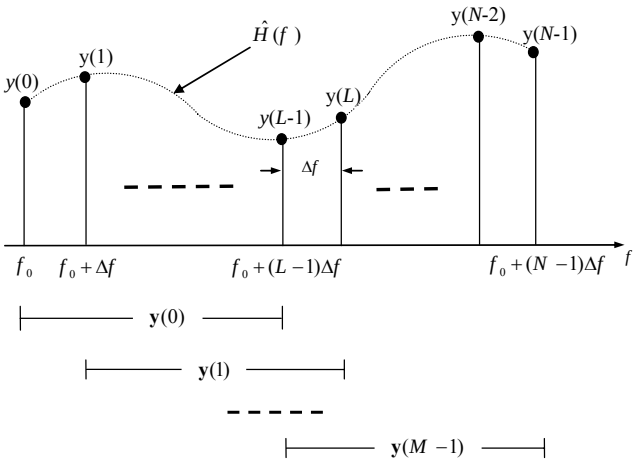

Figure 7. Spatial smoothing preprocessing to improve the estimation of the correlation matrix.

correlation coefficients between α_i' and α_j', i.e., the ith and jth element of the multipath model parameter vector, $\mathbf{a}$. The results of the analysis reveal that the FBCM has a better de-correlation effect than FCM. They also show that the magnitude of the correlation coefficient of the forward-backward estimation method depends on M, Δf, time-delay difference, $\Delta\tau = \left(\tau_i - \tau_j\right)$, length of the sub-array L, phase difference, $\left(\theta_i - \theta_j\right)$, and the lowest frequency f_0. Examples of the graphs corresponding to the correlation coefficients with respect to the number of sub-arrays M and the delay difference $\Delta\tau$ in [4] verifiy that the FBCM has better de-correlation effect than the FCM. Lower correlation coefficients means better performance of the MUSIC algorithms. It is also seen that in both FCM and FBCM with L fixed, an increase in M results in a decrease in the correlation coefficient. With large values of L, the potential for higher resolution of the super-resolution algorithms increases [14]. On the other hand, from Eq. (24), we can see that for a fixed value of N, the value of M decreases as L increases. The decrease in M is equivalent to less smoothing of $\hat{\mathbf{R}}_{\mathbf{yy}}$, resulting in less consistency of the eigenvalues and eigenvectors of $\hat{\mathbf{R}}_{\mathbf{yy}}$. This reduces the number of coherent α_k s that can be detected [13]. Consequently, the value of L needs to be selected so that it provides a balance between resolution and stability of the algorithm. A variety of values for L used in the literature includes $2N/3$ in [13], $3N/4$ in [14], and $3N/5$ in [14]. In this paper, it was experimentally realized that a value of about $2N/5$ led to better performance of the algorithms.

3.4 LFM Signal – Practical Values

The LFM signal (chirp) has been used in traditional acoustic and RF ranging systems for many years [6, 15]. The actual transmitted waveform for range processing in this research is a burst waveform of LFM (chirp) signal. A complex form of an LFM pulse in continuous time interval, $-T_p/2 \le t \le T_p/2$, is given by [16]

$$p(t) = \text{rect}\left(\frac{t}{T_p}\right)\exp\{\text{j}2\pi\varphi_{\text{LFM}}(t)\} \qquad (27)$$

in which

$$\varphi_{LFM}(t) = \frac{1}{2}\frac{B}{T_p}t^2 + f_c t + \varphi_0 \qquad (28)$$

where T_p is the LFM pulse width, $B(=f_2 - f_1)$ is the swept bandwidth, and f_0 is the center frequency. $\gamma = B/T_p$ is called the frequency rate. f_1 and f_2 are called the LFM start frequency and stop frequency, respectively. The derivative of phase determines the instantaneous frequency $f(t)$ of the signal, *i.e.*,

$$f(t) = \frac{d\varphi_{LFM}(t)}{dt} = f_c + \frac{B}{T_p}t \qquad (29)$$

In [16], there is an expression for the Fourier transform of (27) in terms of the Fresnel integrals. Diagrams of $s(t) = \text{Re}\{p(t)\}$, $f(t)$, spectrum of $s(t)$, and its baseband ACF have been depicted in Figure 8(a, b, c, d), respectively. The LFM parameters are: f_1=1*k*Hz, f_2=3*k*Hz, f_c=2*k*Hz, B=2*k*Hz, and $T_p = 50$ ms $(TB = 100)$[1]. It appears from Figure 8(c) that the spectrum would have its most energy in frequency range of $f_c - B/2 \le f \le f_c + B/2$. This is true if TB is large enough. For large TB, the LFM signal has approximately a rectangular spectrum.

The baseband ACF of an LFM pulse at the output of a matched filter, which is matched to an LFM pulse with pulse-width T_p is a very narrow compressed pulse with pulse-width $W \ll T_p$, and some attenuated side-lobes. The width of the compressed pulse W is inversely proportional to the bandwidth B and is given by

$$W = \frac{2}{B} \qquad (30)$$

[1] $TB = T_p \times B$ is called the time-bandwidth product.

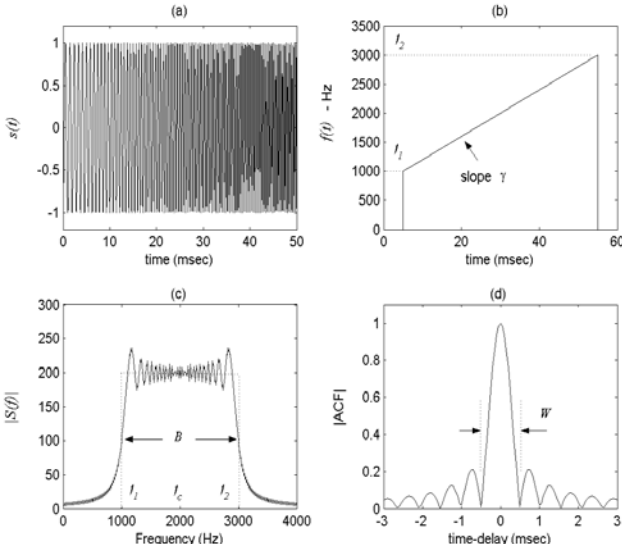

Figure 8. An LFM pulse with TB =100

(a): Time-domain signal, (b): Instantaneous frequency, (c): spectrum, (d): |ACF|

LFM signals are very efficient in suppressing noise and interference [15]. They are spectrally efficient signified by the sharp fall off beyond the $f_c \pm \frac{1}{2}B$ points, *i.e.*, there is little spectral leakage outside the nominal 3dB bandwidth and the spectrum is contiguous and relatively uniform within the 3dB bandwidth. However, they do have some in-band ripples known as Fresnel ripples. The spectrum becomes increasingly rectangular with large TB.

Moreover, they are continuous in time-domain without having many abrupt changes in amplitude or phase. Therefore, the LFM acoustic wave generated by the loudspeaker maintains the signal shape, correlation, and spectral characteristics. Despite many advantages of the LFM signals, on the other hand, constraints such as high background acoustic noise power at low frequencies, signal attenuation due to absorption [3], signal distortion due to the Doppler effect, latency, etc., may not let us choose signal parameters arbitrarily. Table 1 shows optimum empirical values of f_c, B, and T_p. At f_c=2*k*Hz and relative humidity of 30%, the total attenuation due to absorption for a distance of 30m is about 0.7 dB [3]. Besides, with B=2*k*Hz, the spectral interference with the high power portion of the background acoustic noise is negligible, since the background noise has most of its energy in the range of a few hundred Hertz [3]. Also with T_p=50ms, we have TB=100, which means that the signal is spectrally efficient signified by the sharp fall off beyond the $f_c \pm B/2$ points.

3.5 Estimation of $H(f)$

The method of estimating $H(f)$ is the one in which a partial-time or windowed correlation function is first Fourier transformed, and then used to calculate $H(f)$. Following [7], the CCF of a stationary stochastic process is given by

$$r_{xs}(t) = \hat{h}(t) * r_{ss}(t) \tag{31}$$

Taking the Fourier transform of Eq. (31), we obtain

$$\hat{H}(f) = \frac{G_{xs}(f)}{G_{ss}(f)} = \frac{X(f)S^*(f)}{|S(f)|^2} \tag{32}$$

where $G_{ss}(f)$ and $G_{xs}(f)$ are Fourier transforms of $r_{ss}(t)$ and the windowed $r_{xs}(t)$, respectively. Note that, when using Eq. (32), a more accurate estimation of $H(f)$ can be obtained by first choosing $s(t)$ with flat spectrum, e.g. the LFM signal previously discussed, and then dividing over the bandwidth where $|S(f)|$ is sufficiently high. Figure 9 shows the model used for practical implementation of channel estimation. The sampling frequency is 20 kHz, the LFM signal duration T_p=50ms, and the sampling time for each measurement is 150ms. So, the sizes of $s(t)$ and $x(t)$ are 1000 and 3000 samples, respectively. The number of FFT points used is 4096, the next power of 2 over 3000. Following the IFT, the first peak of the cross-correlation function is identified and all data outside of a rectangular window centered at the peak was discarded. Throughout all experiments, it has been realized that the size of the window has been of considerable importance, when isolating the

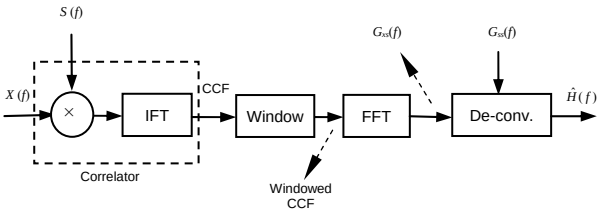

Figure 9. Model implemented in channel estimation

DLOS and its closest multipath components from the rest of the correlation function. It is dependent on the width, W, of the ACF and on the maximum time delay, $\tau_{\max}$, which is used for the calculation of Δf in Eq. (18).

For a bandwidth of 2000 Hz, using Eq. (3.25), the width W of the autocorrelation function main-lobe is about 1ms, so, the number of samples within the main-lobe is about 28 ($\approx 1 \times 4096/150$). If the size of the window is chosen to be 64 samples, then the partial-time correlation function would allow for the multipath components with a maximum delay, $\tau_{\max}$, of about 2.25 ms (referred to the window start time). The partial-time correlation function, is then, Fourier transformed. In order that the FFT results to be more accurate, the number of time-domain samples at the input of the FFT block must be increased. The easiest way to do this is by zero padding [17] of the time-domain signal.

Figure 10 shows examples of $\hat{H}(f)$ for different cases, including the regular and partial-time correlation functions, using rectangular windows with different lengths, w_1 and w_2. In Figure 10, the range corresponding to the DLOS is 370 cm and the transmit LFM signal has a bandwidth of 2 kHz and a duration of 50ms. As the size of the window decreases, the magnitude of $\hat{H}(f)$ becomes smoother and the phase, $\angle\hat{H}(f)$, tends to smaller values; since the number of multipath components confined in the window decreases.

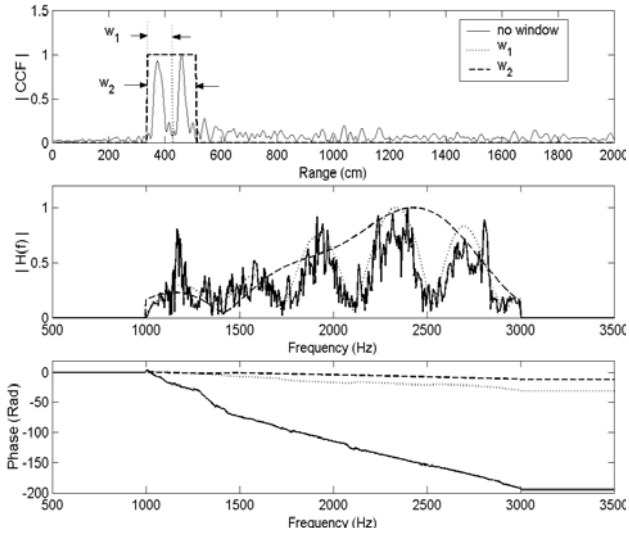

Figure 10. Profiles of $\hat{H}(f)$ at d = 370 cm and effect of rect. windowing .

3.6 MUSIC Parameters – Practical Values of Δf, N, L, M, and L_p

In order to avoid aliasing in the time-domain for MUSIC, the value of Δf must be determined so as to satisfy the condition in Eq. (18). By using the partial-time correlation function, the time interval between 0 and t_d, shown in Figure 11, does not enter the MUSIC algorithms. As a result, the new

maximum time delay, $\tau_{\max}$, corresponding to the latest multipath, now reduces to a lower value confined in the window. Therefore, the value of Δf is now determined by the size of the window. Substituting $\tau_{\max} = 2.25$ ms in Eq. (18), the frequency sampling interval was $\Delta f \leq 222$ Hz. In practice, a value of 70 Hz for Δf also allowed for the generation of the FBCM in Eq. (25) with a reasonable size, L.

The number, N, of the frequency-domain samples required for MUSIC for one measurement snapshot should be equal to the ratio of the signal bandwidth, B, and the frequency sampling interval, Δf, *i.e.*, $N = B/\Delta f$. If $\Delta f = 70$ Hz and $B = 2$kHz, then $N \approx 28$. In the determination of L, it was shown in Chapter 4, Sec. 4.6.2 that to provide a balance between resolution and consistency of MUSIC, different values of L have been used in the literature, including: $N/2$, $N/3$, $3N/4$, and $3N/5$. In this thesis it was realized that a value of 11, which is between $N/3$ and $N/2$, resulted in better performance in terms of accuracy and resolution. With $N = 11$, the estimated FBCM, defined in Eq. (4.23), has dimensions, 11×11. By using Eq. (24), the value of M is calculated to be 18.

If the noise is white and an unlimited number of data samples are available, the $L - L_p$ smallest eigenvalues of the correlation matrix, defined are all equal to σ_w^2. So, the parameter, L_p, the number of signals components, can be easily determined by inspection through the L eigenvalues of the estimated correlation matrix. In practice, however, the noise is not white and the number of data samples is limited. As a result, the noise eigenvalues are not all equal. In [13] there are complicated expressions for the estimation of L_p. Here, we describe the effects of underestimation and overestimation of L_p on the operation of MUSIC. If the estimated value of L_p is less than its true value, the MUSIC algorithms may not cover all the multipath components present, whereas if it is greater than the true value, spurious peaks may appear in the pseudospectrum [4] or zeros on the unit circle. In the experiments, however, the overestimation of L_p did not make any variation in the estimation of the TOA or multipath resolution. The value used for L_p in all experiments was 4.

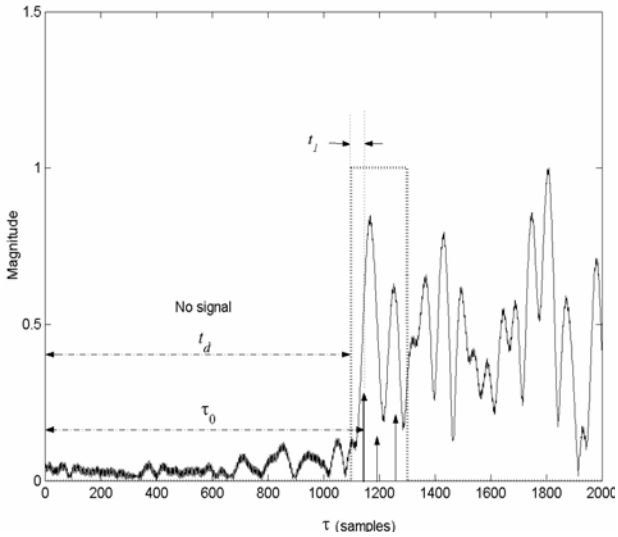

Figure 11. Partial-time Correlation of the MF output by Widowing

Table 1. LFM signal and root-MUSIC parameters

f_c (*k*Hz)	B (*k*Hz)	Δf (Hz)	T_p (ms)	N	L	M	L_p
2	2	70	50	28	11	18	4

4 Performance Evaluation

When the channel is single-path, the first peak of the MF can be used to estimate the TOA and the variance of the estimate is bounded by the Cramer-Rao Lower Bound (CRLB). In multipath channels, however, the CRLB [2] cannot be used and the performance evaluation is accomplished by using empirical measurement data. Due to lack of space, some results have been presented. The overall

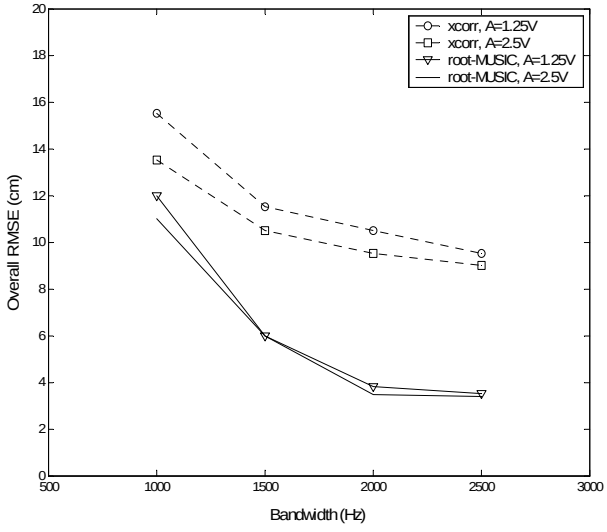

Figure 12. RMSE value of range

RMSE value of the range, which is the RMS range error around its true value, vs. actual range and signal bandwidth is depicted in Figure 12. In this figure, xcorr stands for the cross-correlation method and A is the chirp amplitude. As seen from Figure 12, by using the root-MUSIC algorithm and the same signal, the range accuracy is improved with a factor of at least three. Also, for some ranges, the ability of the system to resolve the DLOS from other multipath components is shown in Figure 13.

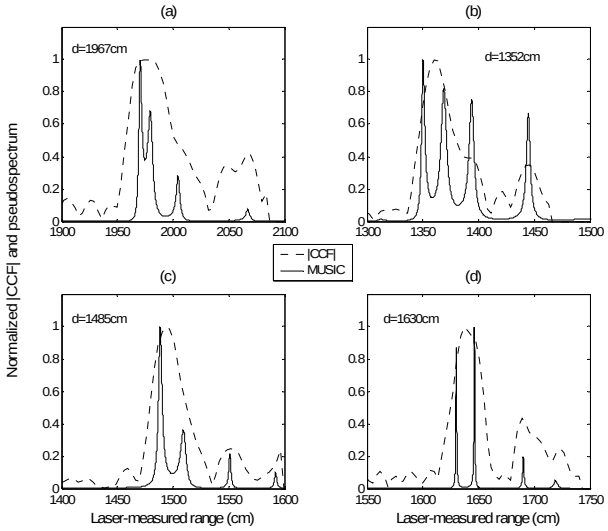

Figure 13. Delay profiles and MUSIC pseudo-spectrum

5 Conclusions

Due to complexity and site-specific nature of indoor localization, a precise model for the performance evaluation cannot be found in the literature and, in this paper, the evaluation has been accomplished based on empirical measurement data. Analysis of the time requirement for the root-MUSIC algorithm is beyond the scope of this paper, but it should be said that the processing time is mostly determined by the size, L, of the correlation matrix, which is 11 in this paper. This value of L is reached by using a partial-time (windowed) CCF containing, specifically, the informative part of the CCF generated by a matched-filter.

Analysis of the empirical data using MATLAB shows that for a signal bandwidth of 2000Hz, the root-MUSIC algorithm reduces the averaged range error bias significantly. It also shows that the accuracy is improved by using a larger bandwidth, since a larger bandwidth can improve the statistical performance of the TOA estimation. The overall performance of the correlation technique in terms of an averaged root-mean-squared-error of range reduces with the root-MUSIC technique by a reduction factor of about $1/3$. It is also shown that when the signal bandwidth is more than 2500Hz, no improvement is observed.

References:

[1] A. Mandal, C. V. Lopes, T. Givargis, A. Haghighat, R. Jurdak, P. Baldi, "Beep: 3D Indoor Positioning Using Audible Sound," *Proceeding of the IEEE Consumer Communications and Networking Conference (CCNC'05)*, Las Vegas, January 2005.

[2] X. Li, K. Pahlavan, "Super-resolution TOA estimation with diversity for indoor geolocation," *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 224-234, Jan. 2004.

[3] ISO. Standard 9613-1: Acoustics - Attenuation of sound during propagation outdoors (part 1), 1993. http://www.measure.demon.co.uk/Acoustics Software/iso9613.html.

[4] H. Yamada, M. Ohmiya, Y. Ogawa, K. Itoh, "Superresolution Techniques for Time-Domain Measurements with Network Analyzer," *IEEE Trans. AP*, vol. 39, no. 2, pp. 177-183, Feb. 1991.

[5] M. A. Hassan, M. R. Azimi-Sadjadi, G. J. Dobeck, "Separation of Multiple Time Delays Using New Spectral Estimation Schemes," *IEEE Trans. SP*, vol. 46, no. 6, June 1998.

[6] A. J. Barabell, "Improving the resolution performance of eigenstructure-based direction-finding algorithms," in *Proc. ICASSP*, Boston, MA, 1983, pp. 336-339.

[7] J. G. Proakis, *Digital Communications*, Third edition, McGraw-Hill, New York, 1995.

[8] B. D. Rao, K. V. S. Hari, "Performance analysis of root-MUSIC," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-37, pp. 1789-1794, 1989.

[9] T. J. Shan, M. Wax, T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806-811, Aug. 1985.

[10] M. N. Vallidis, *WHISPER: A Spread Spectrum Approach to Occlusion in Acoustic Tracking*, PhD

dissertation. University of North Carolina at Chapel Hill, 2002

[11] Marie-Agnes Pallas, "Active High Resolution Time Delay Estimation for Large BT Signals," *IEEE Trans. SP*, vol. 39, no. 4, April 1991.

[12] Simon Haykin, *Adaptive Filter Theory*, Prentice-Hall, Inc., Englewood Cliffs, New Jersy 07632, 1991.

[13] X. Li, *Super-Resolution TOA Estimation with Diversity Techniques for Indoor Geolocation Applications*, PhD dissertation, Worcester Polytechnique Institute, April 2003.

[14] D. Tufts and R. Kumaresan, "Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood," *Proc. of the IEEE*, vol. 70, no. 9, pp. 975-989, Sep. 1982.

[15] G. Cook, A. Zaknich, "Chirp Sounding the Shallow Water Acoustic Channel," *International Conference on Acoustics, Speech and Signal Processing*, Seattle, Washington, USA, IEEE, 4: pp. 2521-2524, 1998.

[16] T. X. Misaridis, J. A. Jensen, "An effective coded excitation scheme based on a predistorted FM signal and an optimized digital filter," *Proc. of IEEE Ultrasonic Symposium*, vol. 2, pp. 1589-1593, 1999.

[17] A. V. Oppenheim, R. W. Schafer, *Discrete Time Signal Processing*, Prentice-Hall Inc., 1989.