TOA estimation of chirp signal in dense multipath environment for low-cost acoustic ranging

TOA Estimation of chirp signal in Dense Multipath Environment for Low-cost Acoustic Ranging

Lei Zhang, Minlin Chen, Xinheng Wang, *Senior Member, IEEE,* and Zhi Wang, *Member, IEEE*

***Abstract*—In this paper, a novel time of arrival (TOA) estimation method is proposed based on an iterative cleaning process to extract the first path signal. The purpose is to address the challenge in dense multipath indoor environments that the power of the first path component is normally smaller than other multipath components, where the traditional match-filtering (MF) based TOA estimator causes huge errors. Along with parameter estimation, the proposed process is trying to detect and extract the first path component by eliminating the strongest multipath component using a band-elimination filter in fractional Fourier Domain (FrFD) at each iterative procedure. To further improve the stability, a slack threshold and a strict threshold are introduced. Six simple and easily calculated termination criteria are proposed to monitor the iterative process. When the iterative 'cleaning' process is done, the outputs include the enhanced first path component and its estimated parameters. Based on these outputs, an optimal reference signal for the match-filtering (MF) estimator can be constructed, and a more accurate TOA estimation can be conveniently obtained. The results from numerical simulations and experimental investigations verified that, for acoustic chirp signal TOA estimation, the accuracy of the proposed method is superior to those obtained by the conventional MF estimators.**

***Index Terms*—TOA estimation, acoustic ranging, multipath, iterative 'cleaning' process.**

I. Introduction

With wide use of smartphones and high demand of indoor location-based services, various approaches for indoor positioning have been proposed by using the technologies based on sound, GSM, Bluetooth, Wi-Fi, light and magnetic fields [1], [2], [3], [4], [5]. Among these approaches, sound-based positioning technology has attracted researchers' attention [6] due to its advantages of being fully compatible with commercial off-the-shelf (COTS) smartphones, relatively higher positioning accuracy and low-cost infrastructure. Low-cost acoustic ranging technology is the base-stone of this kind of indoor localization and navigation systems. Most of the prototype systems are using time of arrival (TOA) estimation of an acoustic chirp signal to realize ranging based positioning [1], [7], [8], [9], [10], because TOA is apparently the most widely used ranging technique in ranging systems [11].

The TOA estimation techniques have been discussed extensively in the past few decades within radar and impulse radio Ultra-Wideband (IR-UWB) systems, and many methods have been proposed, including match-filtering (MF) based coherent methods [14], energy-detection based non-coherent methods [13], and super-resolution estimation method [19]. Benefiting from low-rate sampling, non-coherent methods can be quickly converged [13], but the estimation accuracy is not satisfactory. Coherent methods based on MF can obtain a more robust performance. In [14], the peak of the MF output is considered as the location of the direct path, but this is only applicable for single-path channels with additive white Gaussian noise (AWGN) [15]. For a multipath channel, a threshold detection method is studied and proposed in [16], [18], and [17], in order to detect the direct path based on MF outputs. Different from non-coherent and coherent methods, the super-resolution method is used to detect the first path component in frequency domain, where a multiple signal classification (MUSIC) and independent component analysis (ICA) based method are proposed in [19] and [21], respectively. In [20], a MUSIC based super-resolution method to realize TOA estimation of acoustic chirps is proposed. However, it still needs to firstly estimate the channel impulse response (CIR) using the MF method. Meanwhile, the frequency domain techniques only improve the TOA resolution in flat fading channels which are not feasible assumptions in many applications, and the performance seriously degrades when there are a number of highly correlated interference signals in the same range, especially in lower signal to ratio (SNR) situations [21].

Due to the similarities between acoustic ranging, radar systems and IR-UWB, a fixed-threshold based MF method is widely used to realize TOA estimation of acoustic chirp signals, based on an assumption of full prior knowledge of the source signals [1], [7], [8], [9], [10], where the TOA value is considered as the time delay of the first path component, which is the direct path in line-of-sight (LOS) scenarios. Unfortunately, this assumption of full prior knowledge of source signals is unrealistic in practical applications, especially in low-cost acoustic ranging systems due to the poor performance of electronic components. Accurate TOA estimation in dense multipath environments via smartphones and low-cost infrastructure is still an open problem and many challenges still need to be overcome:

Firstly, indoor multipath propagation is a well-known challenge to the TOA estimation [22]. Compared with radio signals, indoor acoustic signals usually endure a denser multipath propagation due to their poor penetrating capacity. In the geometrical room acoustic theory, the received acoustic signal consists of multiple attenuated and delayed replicas of the source signal due to the acoustic reflections and diffusions [23]. This multipath propagation effect introduces echoes and spectral distortions into the observed signal which severely

L. Zhang, M. Chen and Z. Wang are with State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China (e-mail: zhlei0202@163.com; chenminlin1995@foxmail.com; wangzhizju@gmail.com;)

X. Wang is with the School of Computing and Engineering, University of West London, UK. (e-mail: xinheng.wang@uwl.ac.uk)

deteriorates the source signal [24]. In this case, it is common that the first path component is not the strongest, especially in large space environments. TOA estimation is the detection of the direct path instead of maximum MF output, otherwise, a substantial ranging error occurs. Then, the first path detection strategy of an MF estimator is a key to the accurate TOA estimation.

Secondly, sampling frequency offset between the source signal, transmitter, and receiver is another great challenge. Acoustic ranging systems designed for smartphones are usually used on smart mobile devices and low-cost infrastructure, this makes the full prior knowledge of the transmitted source signals also very hard to be obtained. It is understood that the low-cost acoustic related COTS modules are used for communications and entertainment. Once these modules are used as sensors for ranging measurement, many defects will be exposed. Except for the poor performance and non-consistency of micro-electromechanical systems (MEMS) microphones and speakers, the speed of the crystal oscillator, which provides the clock of the audio sampling and broadcasting system, is usually unstable. This could induce severe signal distortions and frequency shift. For example: a discrete single-frequency signal $s[n]$ is designed under sampling frequency f_s, and the center frequency is f_c; if $s[n]$ is broadcast under an actual sampling frequency f_{st}, the center frequency of the received signal will shift to $f_c \cdot (f_{st}/f_s)$. If we still use the prior knowledge of f_c to realize the TOA estimation, a considerable error will definitely occur.

Thirdly, the frequency shift caused by Doppler Effect is also a challenge, because smart mobile devices are usually carried by human beings. The arbitrary movement of a human being coupled with arm swinging makes the smart mobile devices an extremely complex manoeuvring movement with a high speed. It could introduce an obvious frequency shift of the received signals even at a slow walking speed, due to the low speed of sound propagation.

In a summary, indoor multipath propagation, sampling frequency offset, and Doppler Effect severely degrade the TOA estimation and pose great challenges for accurate acoustic ranging. These problems are strongly needed to be resolved to pave the way for the applications of sound-based indoor localization in the real world. Therefore, aiming to realize a robust TOA estimation via low-cost infrastructure in a dense multipath environment for acoustic ranging, a novel TOA estimation technique based on an iterative 'cleaning' process is proposed. The basic idea of this method is trying to detect and extract the first path component and estimating its TOA using an MF estimator with an optimal reference signal by detecting the maximum of the MF outputs. In this way, the performance of MF based TOA estimator can be greatly improved in dense multipath environments.

The main contributions of this paper are as follows:

- An iterative 'cleaning' process is proposed using a band-pass filter and a band-elimination filter in fractional Fourier Domain (FrFD). The first path component can be detected and extracted from the received signals in FrFD by eliminating the strongest multipath component at each iterative procedure. Along with parameter estimation of the filtered signal during the iterative 'cleaning' process, the influence of sampling frequency offset and Doppler Effect to accurate TOA estimation are greatly mitigated.
- A slack threshold and a strict threshold are proposed in FrFD to dynamically determine a target area which contains the first path component at each iterative procedure, where the first path component is detected, locked and extracted. The strict threshold is dependent on the noise energy, and the scaling factor of the slack threshold is increased with the number of iterations. Through this strategy, the stability of the first path detection can be greatly improved.
- Six termination criteria, which include four successful detection criteria and two abnormal exit criteria, are proposed to guarantee that the iterative 'cleaning' process can be terminated at right time and return the right outputs. All the criteria are simple judgement of logic relations which make it very convenient to use in practical applications.

The remainder of this paper is organized as follows. In Section II, we discuss the problems in TOA estimation of acoustic chirps, and formulate the consequences of unreliable prior knowledge and first path detection strategy to MF based TOA estimator. The details of our proposed novel TOA estimation method is introduced in Section III, including the calculation steps of the iterative 'cleaning' process and six termination criteria. In Section IV, simulations are run to present the detailed process of the proposed method and investigate the performance of TOA estimation and first path detection. Experiments and results are introduced in Section V to demonstrate the applications of the proposed method in a underground car park. At last, we draw our conclusions in Section VI.

II. Problem Formulation

With regard to the TOA estimation based on the MF output for the low-cost acoustic ranging, two main influential factors have severely limited its accuracy and practical applications, which are the unreliable prior knowledge of the transmitted signals and unreliable detection strategy of the first path component.

A. Unreliable prior knowledge of the transmitted signals

The complex formation of a chirp signal is expressed as

$$s(t) = e^{j2\pi(f_0 t + \frac{1}{2}k_0 t^2)}, t \in [0, T], \tag{1}$$

where f_0 and k_0 are the initial frequency and modulation rate, respectively, and T is the time duration of the chirp signal. However, these prior knowledge of source signal becomes no longer reliable when we consider the relative moving speed v between the transmitter and receiver, and the poor performance of COTS modules.

If the frequency of a digital to analog (DA) converter in the transmitter is f_{st} and the sampling frequency at the receiver is

f_{sr}, the initial frequency and modulation rate will change to

$$\begin{cases} f_0^{'} = f_0[\frac{f_{st}}{f_{sr}}(1+\frac{v}{c})] \\ k_0^{'} = k_0[\frac{f_{st}}{f_{sr}}(1+\frac{v}{c})]^2 \\ T^{'} = T/[\frac{f_{st}}{f_{sr}}(1+\frac{v}{c})] \end{cases}, \tag{2}$$

when the received signal is analyzed with prior knowledge of f_s, and c is the propagation speed of the sound. From this relation, we can clearly find that the sampling frequency offset and Doppler Effect can introduce a considerable bias to the initial frequency and modulation rate. Accordingly, the signal with the parameters $f_0^{'}$, $k_0^{'}$ and $T^{'}$ is the actual transmitted signal and denoted as $s^{'}(t)$.

Then, when $s^{'}(t)$ is transmitted over an L *path* fading channel, the complex form of the received signal $x(t)$ can be formulated as

$$x(t) = s^{'}(t) * h(t) = \sum_{i=0}^{L-1} \alpha_i s^{'}(t-\tau_i) + n(t), \tag{3}$$

where $h(t)$ is the channel impulse response (CIR) of room acoustics, α_i and τ_i are the complex channel fading coefficient and propagation delay of the ith propagation path, and $n(t)$ is the additive noise, respectively. Generally, $\tau_0 < \tau_1 < \cdots \tau_{L-1}$, and parameter τ_0 is the TOA of the first path component.

Based on the prior knowledge of source signal to estimated τ_0, that is using $s(t)$ as the reference signal, the MF output $R_{xs}(\tau)$ is

$$\begin{aligned} R_{xs}(\tau) &= \sum_{i=1}^{L} \int_{-\infty}^{+\infty} \alpha_i s^{'}(f) s^*(f) e^{-j2\pi f \tau_i} e^{j2\pi f \tau} df \\ &\quad + \int_{-\infty}^{+\infty} n(f) s^*(f) e^{j2\pi f \tau} df \\ &= \sum_{i=1}^{L} \alpha_i R_{s^{'}s}(\tau) * \delta(\tau - \tau_i) + R_{ns}(\tau), \end{aligned} \tag{4}$$

where $R_{s^{'}s}(\tau)$ is the cross-correlation result of $s^{'}(t)$ and $s(t)$, and $R_{ns}(\tau)$ is the result of noises $n(t)$ and $s(t)$.

In the real world, $n(t)$ usually contains the colored noise. The term $R_{ns}(\tau)$ submerges the weak multipath components and degrades the TOA estimation performance of MF estimator at far distance in the practical use. Except $R_{ns}(\tau)$, from Eq. (4), we can find that the representation of multipath time delay in $R_{xs}(\tau)$ is mainly dependent on $R_{s^{'}s}(\tau)$. The properties of $R_{s^{'}s}(\tau)$ are described as follows:

(1) If $s^{'}(t)$ is identical to $s(t)$ after both energy normalization, which means the prior knowledge is reliable, $R_{s^{'}s}(\tau)$ could be considered as the auto-correlation result. Then $R_{s^{'}s}(\tau) \leq R_{s^{'}s}(0)$. The time delay of each multipath components can be correctly represented by $R_{xs}(\tau)$.

(2) If $s^{'}(t)$ approximates to $s(t)$ after energy normalization, which means the prior knowledge is no longer reliable, $R_{s^{'}s}(\tau) \leq R_{s^{'}s}(\rho)$, where ρ is a constant value which is determined by the difference between $s^{'}(t)$ and $s(t)$. Then, the estimated TOA will be embedded with a bias related with ρ and $R_{n_c s}(\tau)$.

Therefore, if we still use the prior knowledge of initial frequency and modulation rate, which is no longer reliable in the real world, to construct the reference signal for MF estimator, it will introduce a substantial error to TOA estimation. Thus, the parameters of the received signal should be estimated to construct the reference signal for the MF estimator, in order to mitigate the influence of frequency shift caused by sampling frequency offset and Doppler Effect.

B. Unreliable detection strategy of the first path component

Assuming that $s^{'}(t) = s(t)$, at each $\tau = \tau_i$, a positive extremum will definitely appear at the peak envelope of $R_{xs}(\tau)$. Thus, the estimated path delay $\hat{\tau}_i$ can be calculated by

$$\hat{\tau}_i = \underset{\tau}{\text{Extremum}} \left\{ \text{peaks} \left[|R_{xs}(\tau)| \right] \right\}, \tag{5}$$

where $\text{peaks}[\cdot]$ is the peak finding operator, and $\text{Extremum}\{\cdot\}$ is the extremum extraction operator. By using a threshold-based first path detection method, the TOA estimation of the first path component $\hat{\tau}_0$ can be obtained by

$$\hat{\tau}_0 = \min_i \left\{ |R_{xs}(\hat{\tau}_i)| \geq \lambda \max \left[|R_{xs}(\tau)| \right] \right\}, \tag{6}$$

where λ is a scaling factor of the threshold used to detect the first path component. The TOA is considered as the first extremum peak exceeding this threshold. Due to the existence of colored noise, the term $R_{ns}(\tau)$ is not equals to zero, and the value of λ should be carefully chosen.

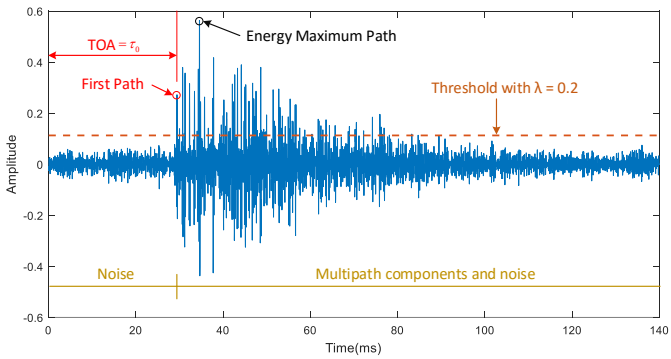


Fig. 1. MF output of received signals in a multipath environment.

Shown in Fig. 1 is the MF output of received signals in a dense multipath environment. Each extreumum can be considered as a signal component received from a propagation path. We can clearly find that the signal component from first path is no longer the one with the strongest power, then, the maximum peak selection strategy for TOA estimation is no longer reliable. When a threshold is chosen appropriately, the first path can be correctly detected and an accurate TOA could be obtained. Otherwise, a substantial TOA estimation bias will occur. If λ is too small, a high false alarm rate or early detection is expected where the peak prior to the first path is detected due to the corruption of colored noise. On the other hand, if λ is too large, there is a high chance of late detection because actual TOA peak is possible below λ due to fading. All of these cases introduce large TOA estimation errors and degrade the performance of acoustic positioning systems dramatically.

Actually, the optimal value of λ is very hard to choose due to the complexity of indoor environment in the real world, because it is mainly depended on the power ratio of the background noise, the first path component and the strongest multipath component. Theoretically estimating a rough range of λ needs more additional prior knowledge, such as the geometrical information of the room and the absorption factors of absorbers. A fast and commonly used approach is experimentally calibrating λ through few ranging tests.

In addition, the unreliable prior knowledge of the source signal further degrades the performance of fixed-threshold based MF estimator. Thus, to achieve a higher ranging accuracy and improve the stability of the MF based TOA estimator, a new technology is needed to mitigate the effects of sampling frequency offset, Doppler Effect, and unreliable detection strategy of the first path component.

III. Proposed Method

In this section, a novel TOA estimation method based on an iterative 'cleaning' process is proposed, aiming to realize a robust TOA estimation via low-cost infrastructure in dense multipath environments for acoustic ranging. The framework of our proposed method is shown in Fig. 2. The received signal $x[n]$ is sent into the iterative 'cleaning' process to detect and extract the first path component in FrFD. At each iterative procedure, the parameters of input signal are firstly estimated to guarantee that all the operations are within the optimal FrFD. The current strongest multipath component is eliminated under the monitoring of termination strategy. When the iterative 'cleaning' process is done, the outputs include the enhanced first path component $x^{'}[n]$ and its estimated parameters $\hat{f}$ and $\hat{k}$. At last, based on these outputs, an optimal reference signal can be constructed for the MF estimator, and a more accurate TOA estimation of acoustic chirp signal can be conveniently obtained in the dense multipath environment for low-cost acoustic ranging.

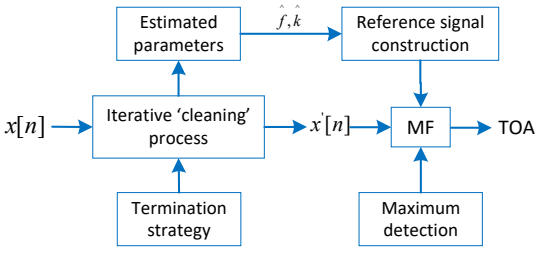


Fig. 2. Diagram of the proposed TOA estimation method.

A. The description of iterative 'cleaning' process

Benefiting from the superior time-frequency representation capacity for the chirp signal [26], fractional Fourier transform (FrFT) technique is chosen to realize the iterative 'cleaning' process. Conventional FrFT based TOA estimation framework, which is proposed in [25], directly extracts the target component using a band-pass filter in the optimal FrFD. Unfortunately, same as the MF estimator, the performance is greatly degraded due to the unreliability of the threshold strategy and the prior knowledge of the source signal.

The block diagram of the proposed iterative 'cleaning' process is shown in Fig. 3. Different from the framework in [25], our proposed iterative 'cleaning' process is mainly using a band-elimination filter to eliminate the current strongest multipath component in the optimal FrFD at each iterative procedure. Compared with directly claiming a component being the first path component in the non-optimal FrFD, it is safer to eliminate the strongest component which is more certain to be a multipath component. Therefore, instead of directly detecting the first path component, we use the iterative 'cleaning' process to conservatively 'wait' for the appearance of the first path component.

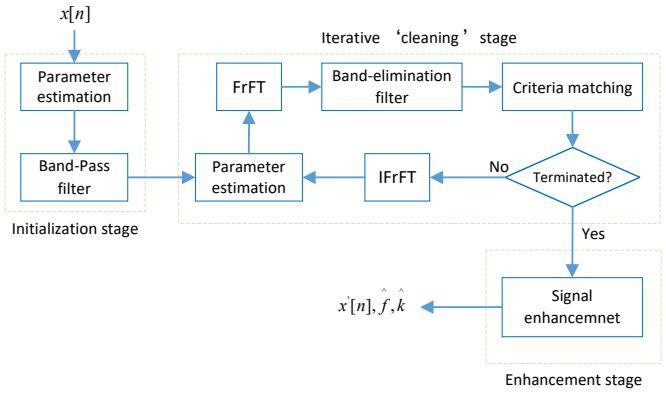


Fig. 3. Diagram of the proposed iterative 'cleaning' process.

It is essential to guarantee that all the operations are within the optimal FrFD, where the strongest components always have the best representation, in order to reduce the energy losing of the weak first path component during the elimination operation. Then, parameter estimation is always the first operation at each iterative procedure to obtain the parameters of the strongest multipath component.

At the ith iterative procedure, the input signal is denoted as $x^i[n]$. The first operation is estimating the parameters of $x^i[n]$, which are the angle of FrFT α, initial frequency $\hat{f}^i$ and modulation rate $\hat{k}^i$. $X_\alpha^i[u]$ is the digital fractional Fourier spectrum of $x^i[n]$ by using FrFT with angle α. By eliminating the strongest multipath component in $X_\alpha^i[u]$, the filtered signal $x^{i+1}[n]$ is obtained by using inverse fractional Fourier transform (IFrFT).

With the progression of the iterative 'cleaning' process, the first path component will gradually become the strongest one. And if any criterion of termination strategy is triggered, the iterative process will be terminated immediately and the filtered signal $x^{'}[n]$ with its parameters are returned.

B. Parameter estimation

Many methods for chirp parameter estimation have been proposed in the literature, including maximum likelihood estimation (ML) [27], and time-frequency analysis based estimation such as multinomial phase parameter estimation [28], Wigner-Hough transform [29], and Radon-ambiguity transform

(RAT) [31]. Computation load is an especially concerned problem in the parameter estimation. In our approach, the ML estimator can be calculated by general cross-correlation to reduce the computation load.

In the presence of AWGN, an ML estimator can achieve the Cramer-Rao lower bound (CRLB) in a single-path channel [30]. This performance can't be achieved in the dense multi-path environment, due to the interference between multipath components. But, the ML estimator still can be used to estimate the parameters of the strongest multipath component.

When the digital time duration of the received signal $x[n]$ is N, based on the ML estimator, the estimated initial frequency $\hat{f}$ and modulation rate $\hat{k}$ are given by finding

$$[\hat{f},\hat{k}] = \max_{f,k,d}|G(f,k)| = \max_{f,k,d}\left|\sum_{n=0}^{N-1} x[n]r^*[n-d]\right|, \quad (7)$$

where $r[n] = e^{j2\pi[fn\Delta+\frac{1}{2}k(n\Delta)^2]}$, $n = 1, 2, \cdots, N^{'}$, is the reference signal. $N^{'}$ is the time duration of $r[n]$, and can be calculated by $N^{'} = \lfloor T^{'}/\Delta \rfloor$, where $\Delta = 1/fs$. To obtain the values of $\hat{f}$ and $\hat{k}$, we need to search a three-dimension parameter space which makes it not being used via low-cost infrastructures.

According to Eq. (2), we can establish the relation between $\hat{f}$ and $\hat{k}$ as

$$\hat{f} = f_0\sqrt{\frac{\hat{k}}{k_0}}. \quad (8)$$

Due to the unknown value of $T^{'}$, we can construct the reference signal $r[n]$ by setting $f = 0$ and $N^{'} = \min[N, \lfloor f_s/2k \rfloor]$ to further reduce the dimension of the searched parameter space. The Eq. (7) can be rewritten as

$$\hat{k} = \max_{k,d}\left|\sum_{n=-N^{'}}^{N+N^{'}-1} x[n]r^*[n-d]\right| = \max_{k,d}|R_{xr}[k,d]|, \quad (9)$$

where $R_{xr}[k,d]$ is the cross-correlation result of $x[n]$ and $r[n]$.

Through Eq. (9), we can scan the whole time-frequency plane to estimate $\hat{k}$ by using a fast Fourier transform (FFT), and the dimension of searched parameter space is reduced to one. In order to further reduce the computation load, modulation rate could be searched over a limited parameter space $|k - k_0| \leq \rho$.

C. Filtering in FrFD

The FrFT of signal $s(t)$ is considered as a rotation in the time-frequency plane with an angle α, which is defined as [32]

$$F^{\alpha}[s(t)] = S_{\alpha}(u) = \int_{-\infty}^{+\infty} s(t)K_{\alpha}(u,t)dt, \quad (10)$$

$$K_{\alpha}(t,u) = \begin{cases} \sqrt{\frac{1-j\cot\alpha}{2\pi}}e^{j(\frac{u^2+t^2}{2}\cot\alpha - ut\csc\alpha)} & \alpha \neq n\pi \\ \delta(u-t) & \alpha = 2n\pi \\ \delta(u+t) & \alpha = (2n-1)\pi \end{cases} \quad (11)$$

where $K_{\alpha}(t,u)$ is the transformation kernel and n is an integer. The signal can be easily recovered from FrFD by using IFrFT, which is the FrFT with angle $-\alpha$:

$$s(t) = F^{-\alpha}[S_{\alpha}(u)] = \int_{-\infty}^{+\infty} S_{\alpha}(u)K_{-\alpha}(u,t)dt. \quad (12)$$

Applying Eq. (10) to Eq. (1), when $k_0 + \cot\alpha = 0$, the Fractional Fourier spectrum $s(t)$ is

$$|S_{\alpha}(u)| = \frac{T}{|\sin\alpha|^{1/2}}\mathrm{sinc}[\pi(f_0 - u\csc\alpha)T], \quad (13)$$

where $\mathrm{sinc}[\cdot]$ represents the sinc function. It is very clear to find that the energy of the chirp signal is concentrated in the bandwidth of $B_m = |2\cdot\sin\alpha/T|$. If and only if $k_0 + \cot\alpha = 0$ and $f_0 - u\csc\alpha = 0$, $|S_{\alpha}(u)|$ reaches the peak value. When signal $s(t)$ is delayed by τ, the FrFT of the delayed signal $s(t-\tau)$ can be written as

$$F^{\alpha}[s(t-\tau)] = S_{\alpha}(u - \tau\cos\alpha)e^{j(\frac{\tau^2\sin\alpha\cos\alpha}{2} - u\tau\sin\alpha)}, \quad (14)$$

which suggests that $|F^{\alpha}[s(t-\tau)]|$ reaches the peak value at $u = \tau\cos\alpha$. Thus, we denote this position in FrFD as u_0 corresponding to τ_0 in MF output.

Then, the delay estimation based on FrFT can behave with respect to the location of the extremum of the Fractional Fourier spectrum. Based on these properties, we can very easily extract or eliminate a signal component in $x(t)$ by multiplying a window function to its fractional Fourier spectrum $X_{\alpha}(u)$ and transform the result into time domain. This process is expressed as

$$x^{'}(t) = F^{-\alpha}[X_{\alpha}(u)w(u)], \quad (15)$$

where $x^{'}(t)$ is the filtered signal, and $w(u)$ is a specially designed window function. The digital form is

$$x^{'}[n] = F^{-\alpha}[X_{\alpha}[u]w[u]]. \quad (16)$$

In this paper, we use the digital FrFT algorithm proposed by Haldun M. Ozaktas and et al in [33] to realize the transform operation between time domain and FrFD.

The optimal rotation angle α can be obtained from the relation $k_0 + \cot\alpha = 0$ and the estimated $\hat{k}$. Due to the dimensional normalization in the digital FrFT algorithm, α is given by

$$\alpha = \begin{cases} \pi + \mathrm{arccot}(-\hat{k}N\Delta^2), & \hat{k} \geq 0 \\ \mathrm{arccot}(-\hat{k}N\Delta^2), & \hat{k} < 0 \end{cases}. \quad (17)$$

For the band-pass filter in FrFD, we propose a window function which is the combination of a Blackman window and a rectangular window, and the digital expression is:

$$w_P[n] = \begin{cases} B[n+G+1], & -G \leq n \leq -1 - N_r \\ 1, & -N_r \leq n \leq N_r \\ B[n+N_b-N_r], & N_r + 1 \leq n \leq G \end{cases}, \quad (18)$$

where $B[n]$ is the Blackman function in time domain, which is

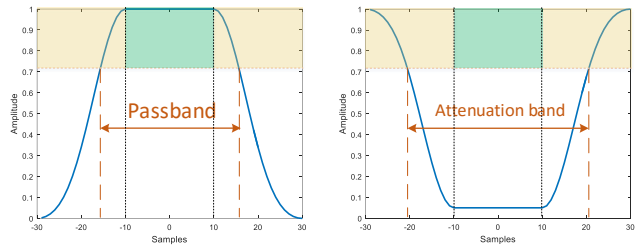
$$B[n] = 0.42 - 0.5\cos(\pi n/N_b) + 0.08\cos(2\pi n/N_b). \quad (19)$$

, N_r is the half-length of the rectangular window or band-pass in FrFD, N_b is the half-length of the Blackman function, and $G = N_b + N_r$. Accordingly, the window function of the band-elimination filter is:

$$w_A[n] = 1 - gw[n], \quad (20)$$

where g is the gain in the attenuation band. Shown in Fig. 4 are the diagrams of $w_P[n]$ and $w_A[n]$ in FrFD with the

parameters of $N_r = 10$, $N_b = 20$ and $g = 0.95$. In order to extract or eliminate one chirp component, these parameters should be chosen carefully to reduce the interference of other components and avoid introducing a 'ghost' signal.



(a) $w_P[n]$ (b) $w_A[n]$

Fig. 4. Diagrams of window function in FrFD.

From Eq. (13), the most energy of a chirp signal is concentrated in the bandwidth of B_m in FrFD. Thus, the value of N_r should equals to B_m to guarantee that the most energy of a chirp signal is extracted or eliminated. Due to the dimensional normalization in the digital FrFT algorithm, the value of N_r can be given as

$$N_r = \left\lceil \frac{1}{2} B_m N \Delta \right\rceil . \tag{21}$$

Generally, we can set $N_b = 2N_r + 1$ for simplicity in practical applications.

D. The algorithm of iterative 'cleaning' process

The successful rate of the first path detection directly impacts the performance of TOA estimation. In order to improve the successful rate, a slack threshold and a strict threshold are proposed in FrFD to dynamically determine a target area which contains the first path component at each iterative procedure. The first path component is detected and extracted from this target area, which is between u_λ^i and u_γ in FrFD. u_λ^i and u_γ are estimated positions obtained by a slack threshold and a strict threshold, where the superscript i denotes the ith iterative procedure. The slack threshold is a criterion same as Eq. (6) and the scaling factor increases with iterative process. The strict threshold is based on the noise power. Generally, $u_\gamma \le u_0 \le u_\lambda^i$.

The detailed process of the 'cleaning' process is shown in Fig. 8, and the main steps are as follows:

Step1: Band-pass filtering and initialization

(1) pre-filter $x[n]$ using an FIR (Finite Impulse Response) digital band-pass filter to eliminate the frequency components out of band $[f_0, k_0 T\Delta]$, and the obtained signal is denoted as $x^0[n]$;

(2) estimate the parameters of $x^0[n]$ using Eq. (8) and Eq. (9), and denote the initially estimated parameters as $\hat{k}^0$, $\hat{f}^0$, where the superscript 0 denotes the initial step of the iterative process;

(3) calculate angle α by Eq. (17) with initially estimated $\hat{k}^0$; calculate FrFT of signal $x^0[n]$ with angle α and denote it as $X_\alpha^0[u]$;

(4) estimate the time delay of the strongest multipath component in $X_\alpha^0[u]$ by using

$$u_{max}^0 = \max_u \left[\left| X_\alpha^0[u] \right| \right] ; \tag{22}$$

estimate the time delay of the first path component in FrFD based on the slack threshold method and the strict threshold method by using

$$\begin{cases} u_p = \text{peaks}\left[\left| X_\alpha^0[u] \right| \right] \\ u_\lambda^0 = \min\limits_{u_p} \left\{ \left| X_\alpha^0[u_p] \right| \ge \lambda_0 \left| X_\alpha^0[u_{max}^0] \right| \right\} \\ u_\gamma = \min\limits_{u_p} \left\{ \left| X_\alpha^0[u_p] \right| \ge \gamma \right\} \end{cases} , \tag{23}$$

where λ_0 denotes the initial value of the scaling factor λ of the slack threshold, and γ is the strict threshold obtained by choosing the maximum value of noise part in $\left| X_\alpha^0[u] \right|$. It is certain that the relationship of u_{max}^0, u_λ^0 and u_γ is $u_\gamma \le u_\lambda^0 \le u_{max}^0$;

(5) design the window function $w[u]$ of band-pass filter as

$$w[u] = \begin{cases} w_P[u - u_{max}^0], & u_{max}^0 - G \le u \le u_{max}^0 + G \\ 0, & others \end{cases} \tag{24}$$

where $G = N_r + N_b$, $N_r = \left| u_\gamma - u_{max}^0 \right| + B_p$ and $N_b \ge \lceil B_m N \Delta \rceil + 1$. B_p is a protection band to ensure that the first path component is contained in the pass band of the designed filter;

(6) use Eq. (16) to filter out a part of multipath components and noise in the signal $x^0[n]$ in FrFD, and recover the result into time domain with $-\alpha$, that is $x^1[n] = F^{-\alpha}[X_\alpha^0[u] w[u]]$;

(7) input $x^1[n]$ into ***Step2***, the iterative 'cleaning' stage.

Step2: Iterative 'cleaning' stage

Assuming that the current iterative time is i, then:

(1) estimate parameters $\hat{k}^i$ and $\hat{f}^i$ of $x^i[n]$; renew α with $\hat{k}^i$; calculate the FrFT of signal $x^i[n]$ with α and denote it as $X_\alpha^i[u]$;

(2) estimate u_{max}^i and u_λ^i by using

$$\begin{cases} u_p = \text{peaks}\left[\left| X_\alpha^i[u_p] \right| \right] \\ u_{\max}^i = \max\limits_u \left[\left| X_\alpha^i[u] \right| \right] \\ u_\lambda^i = \min\limits_{u_p} \left\{ \left| X_\alpha^i[u_p] \right| \ge \lambda_i \left| X_\alpha^i[u_{\max}^i] \right| \right\} \end{cases} , \tag{25}$$

where the scaling factor of the slack threshold $\lambda_i = \lambda_{i-1} + \delta\lambda$, and $\delta\lambda$ is a factor compensation value of λ to improve the stability of first path detection during the iterative process;

(3) estimate the set of target components u_l by finding

$$u_l = \{ \forall u_p | u_p \ge u_{max}^i, \left| X_\alpha^i[u_p] \right| > 0.3 \cdot \left| X_\alpha^i[u_{max}^i] \right| \}, \tag{26}$$

where the elements, denoted as u_{lj}, are the strongest multipath component and the other target components to be eliminated.

(4) update the window function $w[u]$ of band-elimination filter as

$$w[u] = \prod w_j[u],$$

$$w_j[u] = \begin{cases} w_A[u - u_{lj}], & u_{lj} - G \le u \le u_{lj} + G \\ 1, & others \end{cases} , \tag{27}$$

where $G = N_r + N_b$, the value of N_r is obtained by Eq. (21), $N_b \geq 2N_r + 1$ and $g = 0.8$;

(5) use Eq. (16) to filter out the strongest multipath component and the other target components in the signal $x^i[n]$ in FrFD, and recover the result into time domain with α, that is $x^{i+1}[n] = F^{-\alpha}\left[X_\alpha^i[u]w[u]\right]$;

(6) calculate the termination criteria and judge whether the iterative process is completed or not based on the termination strategy; if any termination criterion is satisfied, the iterative process will be terminated immediately. A signal will be sent to ***Step3*** as the input signal based on the *Case* category of termination criteria; otherwise, continue to execute the rest operations.

(7) start the $(i+1)$th iterative procedure with $x^{i+1}[n]$ as the input signal.

Step3: Signal enhancement and result output

Assuming that the input signal is $x^p[n]$, we need to repeat operations (1)∼(5) of ***Step2*** *once again. And the set of target components* u_l *of operation (3)* is given by

$$u_l = \{\forall u_p | u_p < u_{max}^p, |X_\alpha^p[u_p]| \geq 0.3 \cdot |X_\alpha^p[u_{max}^p]|\}, \quad (28)$$

to enhance the filtered signal. The output of the iterative 'cleaning' process includes the filtered signal $x^{'}[n] = x^{p+1}[n]$ and its parameters: $\hat{k} = \hat{k}^p$ and $\hat{f} = \hat{f}^p$.

E. Termination strategy

The termination strategy is essential for the stability of the proposed iterative 'cleaning' process. The proposed termination strategy must guarantee that the iterative process could be terminated at right time. For this purpose, the first approach we thought was using the machine learning to identify the spectrum change at each iterative procedure. However, it usually has to endure heavy computation load. Aiming to make this method being used on COTS mobile devices and low-cost infrastructure, the proposed termination strategy must be simply and easily calculated. Under this condition, the proposed termination strategy includes four successful detection criteria and two abnormal exit criteria. The successful detection criteria are used to avoid 'inadequacy-cleaning', and the abnormal exit criteria are designed for avoiding 'over-cleaning'.

All the criteria are simple judgement of logic relations which make them very convenient to use in practical applications. The iterative process will be terminated when any criterion is triggered.

1) Successful detection criteria: There are four conditions being considered as successful detection of the first path component, as shown in Fig. 5, and listed as follow:

- case 1: $u_{max}^i = u_\lambda^i = u_\gamma$, the iterative process is terminated at ith iterative procedure, and the input signal of ***Step3*** *is* $x^i[n]$;
- case 2: $u_{max}^i = u_\gamma$ and $u_\gamma - u_\lambda^i \leq \delta u$, the iterative process is terminated at ith iterative procedure, and the input signal of ***Step3*** *is* $x^{i+1}[n]$;
- case 3: $u_{max}^i = u_\gamma$ and $u_\gamma - u_\lambda^i > \delta u$, the iterative process is terminated at ith iterative procedure, and the input signal of ***Step3*** *is* $x^i[n]$;
- case 4: $u_{max}^i = u_\lambda^i > u_\gamma$, the iterative process is terminated at ith iterative procedure, and the input signal of ***Step3*** *is* $x^i[n]$.

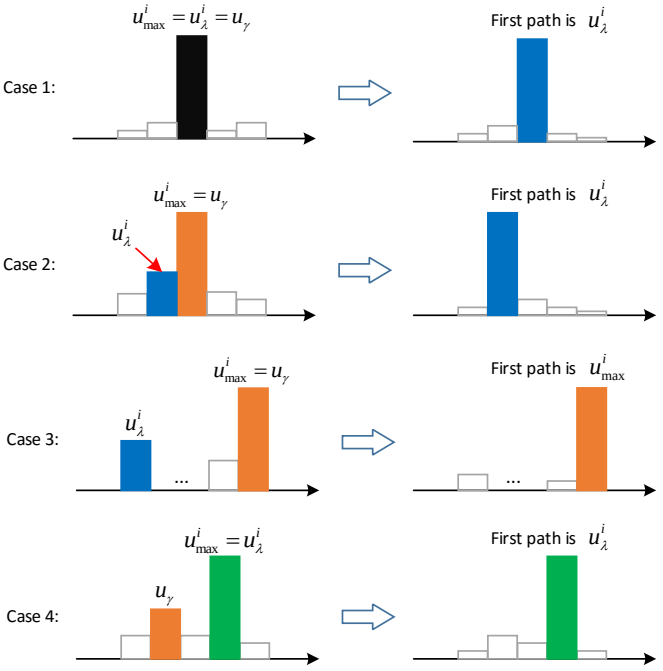


Fig. 5. Termination conditions of iterative process and detected first path.

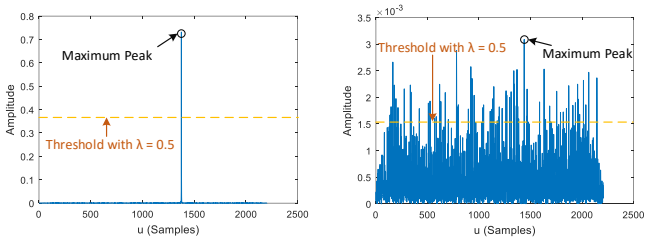
Compared with u_γ, we give a higher confidence level to u_λ^i. $Case1$ and $case4$ are the two normal results when the strongest component appears in the target area. $Case2$ and $case3$ are the two special conditions that happened when the initially estimated parameters have a substantial bias which makes the spectrum in FrFD unable to correctly express the energy and delay property of chirp components. This will greatly increase the probability of a condition that the first path component falls outside of the target area $[u_\gamma, u_\lambda^i]$. By eliminating the components with the strongest power and estimating modulation rate $\hat{k}^i$ during the iterative process, the first path will appear and could be detected by setting a safe distance δu.

2) Abnormal exit criteria: To further improve the stability of the proposed iterative process, abnormal exit criteria are necessary to break the iterative process when the first path is falsely eliminated, by monitoring the spectrum change in FrFD at real time. Thus, we first compare the differences of spectrum in FrFD before and after eliminating the correct first path component. Those differences can be characterized as the criteria to break the iterative process.

Shown in Fig. 6 is the comparison of the fractional Fourier spectrum before and after eliminating a main component of a received single $x[n]$ in FrFD. $x[n]$ is a mono-component signal and expressed as $x[n] = s[n] + N[n]$. The parameters of source signal $s[n]$ are: $f_0 = 3\,\text{kHz}$, $k_0 = 100\,\text{kHz/s}$, $T = 50\,\text{ms}$ and $f_s = 44.1\,\text{kHz}$. The noise $N[n]$ is additive Gaussian noise and the SNR is $0\,\text{dB}$. The fractional Fourier spectrum of $x[n]$ is calculated with optimal rotation angle $\alpha = 0.536 \cdot \pi$ and shown in Fig. 6(a). It is clear that the energy of $s[n]$ is well

concentrated in a quit narrow band and the noise spectrum is well suppressed in FrFD.

Eliminating component $s[n]$ by the band-elimination filter based on the window function of Eq. (27) with $N_r = 1$ and $N_b = 3$, the result is shown in Fig. 6(b). There is a big difference before and after eliminating $s[n]$ from $x[n]$ in FrFD. Based on this change, we can characterize the abnormal exit criteria.



(a) Spectrum of received signal $x[n]$ (b) Spectrum after eliminating $s[n]$

Fig. 6. Contrast of $|X_\alpha[u]|$ before and after eliminating $s[n]$.

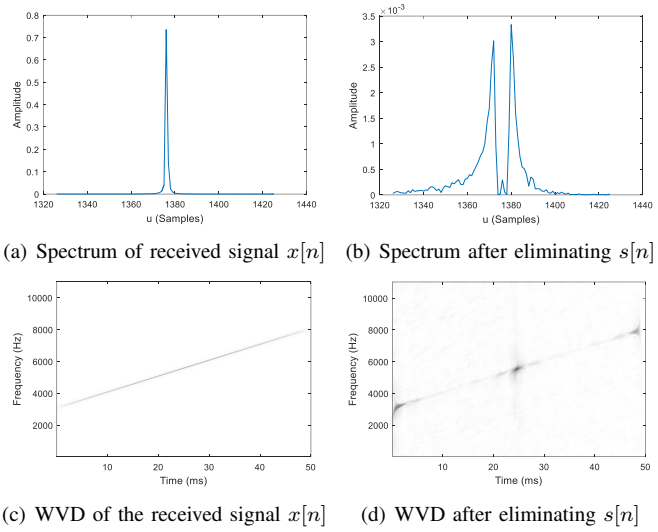
One exit criterion is to set up a bound to the number of peaks exceeding a peak power threshold which is the same as Eq. (6). This variable is denoted as n_p^i at ith iterative procedure. To improve the stability, another supplementary criterion is proposed by using the ratio of peak number at ith and $(i-1)$th iterative procedure which is expressed as $r^i = n_p^{i-1}/n_p^i$. Then, the abnormal exit criteria are:

- if $n_p^i > N_p$, the iterative process is terminated at ith iterative procedure, and the input signal of ***Step3*** is $x^{i-1}[n]$;
- if $r^i > 4$, the iterative process is terminated at ith iterative procedure, and the input signal of ***Step3*** is $x^{i-1}[n]$;

where N_p is a threshold of peak number.

In practical applications, N_p can be set to 15 from experience, and $n_p^0 = 1000$ to avoid iterative process being terminated at first cycle. The value selection of threshold r^i, which equals to 4, should consider the 'ghost' phenomenon after eliminating a component at a high SNR. Shown in Fig. 7 is the 'ghost' phenomenon when the first path component is eliminated by a band-elimination filter with N_r and N_b obtained by Eq. (21) under SNR $= 20$ dB.

It should be noted that 'ghost' phenomenon only appears at high SNRs, and the cause is the rudimental energy of target component. Because the band-elimination filter with parameters obtained by Eq. (21) only eliminates the most power of the target component, the rudimental energy is still large enough to make it observed at the time-frequency plane. If we extend the value of N_r and N_b, the 'ghost' of the target component can be completely killed. However, based on the projection property of FrFT, the bandwidth of N_r in time domain is given by $B_r = N_r \sec(\alpha)$. It is very clear that the oversized value of N_r and N_b can greatly decrease the time resolution of the filter. Actually, the proposed N_r and N_b from Eq. (21) can be considered as a trade-off between high time resolution and filtering performance. When the SNR is low, this 'ghost' is submerged by the noise components.



(a) Spectrum of received signal $x[n]$ (b) Spectrum after eliminating $s[n]$

(c) WVD of the received signal $x[n]$ (d) WVD after eliminating $s[n]$

Fig. 7. Ghost phenomenon of band-elimination filter when parameters are chosen as $N_r = \lceil \frac{1}{2} B_m N \Delta \rceil$ and $N_b = 2N_r + 1$ at high SNR.

IV. Numerical Simulation

Simulations have been run on Matlab to demonstrate the detailed process of the proposed method, and, investigate the performance of TOA estimation and first path detection. We will firstly use a simple indoor propagation model to demonstrate the detailed process of the proposed method. Then, the performance of TOA estimation and successful rate for first path detection is evaluated based on an indoor CIR generator. The following setting has been chosen. The formation of source signal $s[n]$ is chosen as the digital formation of Eq. (1). The simulation parameters and applied values are shown in Table I.

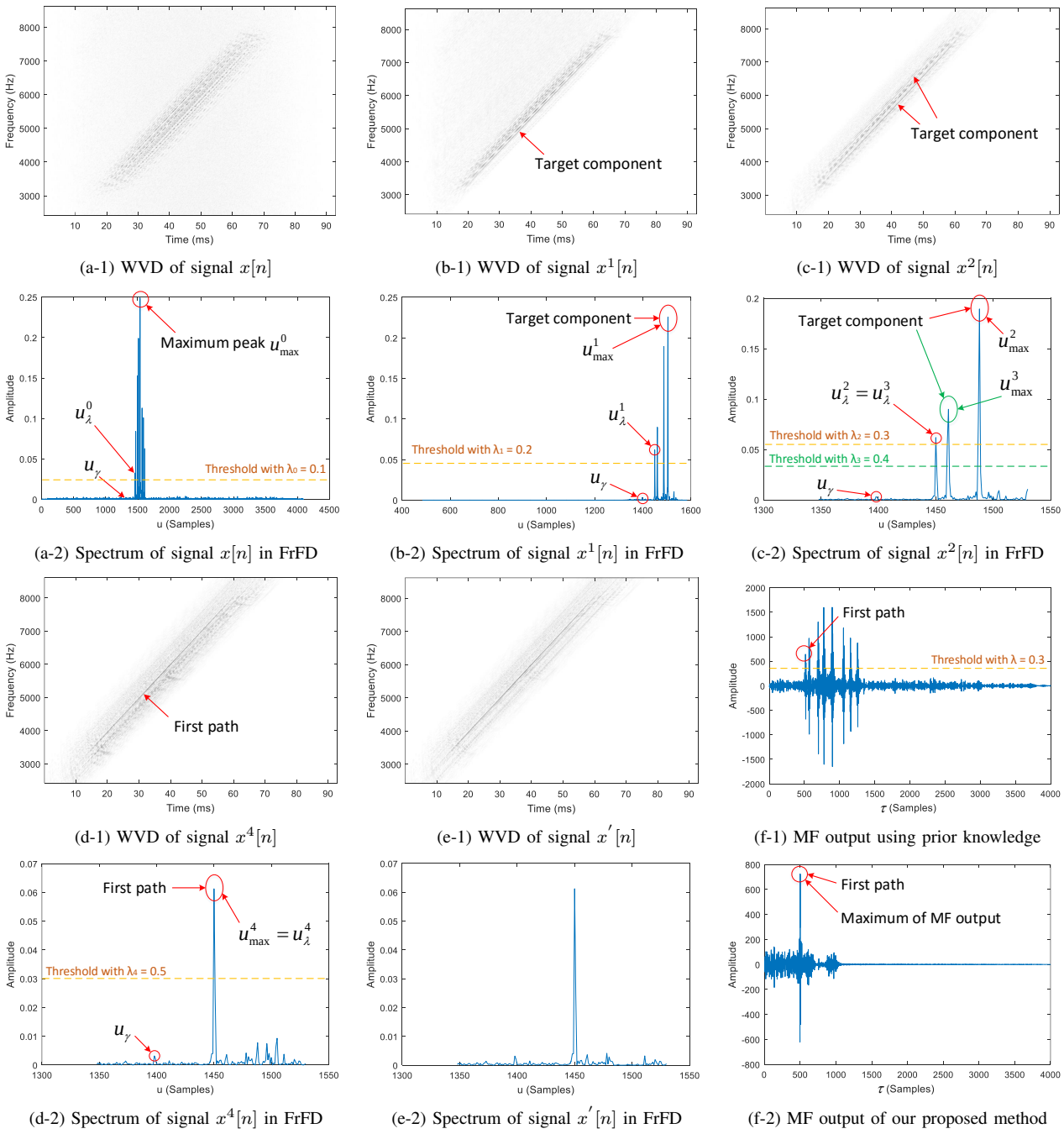
TABLE I
Simulation parameters and applied values.

Objects	Parameter	Definition	Value
Source signal $s[n]$	f_0	Initial frequency	3 kHz
	k_0	Modulation frequency	100 kHz/s
	T	Time duration	50 ms
	f_s	Sampling frequency	44.1 kHz
Transmitter	f_{st}	DA converter frequency	44.3 kHz
Receiver	f_{sr}	Sampling frequency	44.1 kHz
	v	Moving speed	1 m/s

A. The detailed process of the proposed method

Assuming that an indoor channel has eight propagation paths with AWGN noise, and SNR $= 0$ dB. The parameter sets of CIR are $\{\alpha_i\} = \{0.4, 0.6, 0.75, 0.9, 0.95, 0.7, 0.6, 0.5\}$ and $\{\tau_i/\Delta\} = \{500, 550, 680, 760, 880, 1040, 1140, 1240\}$, respectively, where $i = 0, 1, \cdots, 7$ and $\Delta = 1/f_s$. The moving speed of a receiver v is a relative speed corresponding to a transmitter which equals to 1 m/s. Accordingly, the parameters of the actual transmitted signal $s'[n]$ are $f_0' = 3022.5$ Hz, $k_0' = 101\,504$ Hz/s and $T' = 49.6$ ms. The observation time of the received signal $x[n]$ is $N = 4096$.

Shown in Fig. 8 is the calculating process of the proposed method. We also use time frequency distribution WVD to show the details of signal change at each step. The WVD and



(a-1) WVD of signal $x[n]$ (b-1) WVD of signal $x^1[n]$ (c-1) WVD of signal $x^2[n]$

(a-2) Spectrum of signal $x[n]$ in FrFD (b-2) Spectrum of signal $x^1[n]$ in FrFD (c-2) Spectrum of signal $x^2[n]$ in FrFD

(d-1) WVD of signal $x^4[n]$ (e-1) WVD of signal $x^{'}[n]$ (f-1) MF output using prior knowledge

(d-2) Spectrum of signal $x^4[n]$ in FrFD (e-2) Spectrum of signal $x^{'}[n]$ in FrFD (f-2) MF output of our proposed method

Fig. 8. Calculating process of our proposed method.

spectrum in FrFD of the received signal $x[n]$ are shown in Fig. 8(a-1) and Fig. 8(a-2), respectively. Based on ***Step1*** of the proposed iterative 'cleaning' process, the last four components with $\{\tau_i/\Delta\} = \{880, 1040, 1140, 1240\}$ are eliminated, including the strongest power component. The filtered signals are shown in Fig. 8(b-1) and Fig. 8(b-2). The parameters used in ***Step1*** are chosen as $B = 1000$ and $N_b = 50$. Figures from Fig. 8(c) to Fig. 8(d) are obtained by ***Step2*** with 3 iterative procedures. Multiple components are gradually eliminated from τ_3 to τ_1. At 4th iterative procedure, the process is terminated by the successful detection criterion $case1$. Then, the recovered signal is input into ***Step3*** to further filter out noise and enhance the detected signal component. In the end, the output of iterative 'cleaning' process is the signal $x^{'}[n]$ shown in Fig. 8(e-1) and Fig. 8(e-2) with $\hat{k} = 101\,526$ Hz/s. The values of parameters during each iterative procedure are shown in Table II.

Based on the MF output of $x[n]$ and $s[n]$, as shown in Fig. 8(f-1), the estimated TOA is $\hat{\tau}_0 = 476\Delta$ based on the threshold method where $\lambda = 0.3$. Fig. 8(f-2) is the MF output of $x^{'}[n]$ and designed reference signal $r[n]$ with $\hat{k} = 101\,526$ Hz/s and $\hat{f} = 3022.8$ Hz. Using detection of maximum MF output based on the reference signal constructed with estimated parameters, the estimated TOA is $\hat{\tau}_0 = 503\Delta$. Comparing the two TOA estimation methods, the proposed method can achieve a more

TABLE II
THE VALUE OF PROCEDURE PARAMETERS.

iterative time	λ_i	$\hat{k}^i$	u^i_{max}	u^i_{λ}	n^i_p	r^i
0 ($step1$)	0.1	101526	1530	1450	100	–
1	0.2	101518	1505	1450	2	0.02
2	0.3	101836	1488	1450	2	1
3	0.4	101526	1461	1450	2	1
4	0.5	101526	1450	1450	1	0.5
5 ($step3$)	0.3	101526	1450	1450	–	–

accurate result.

B. Performance investigation

The following simulation is designed to investigate the performance of TOA estimation and the successful rate of first path detection under dense multipath environments. The root mean square error (RMSE) is used to analyze the statistical performance of our method and conventional fixed-threshold based MF estimator. A room impulse response generator proposed in [34] is used to simulate the dense multipath propagations. The environment is chosen as a 3-dimensional room with $40 \times 3 \times 2.7$(m). The positions of the receiver and transmitter are at $[2, 1.5, 1]$(m) and $[2 + D, 1.5, 1]$(m) respectively. The variable $D = 1, 2, \cdots, 30$ (m) is the distance between the transmitter and the receiver. At the same time, the reflection coefficient and order are chosen as 0.5 and -1. The simulation at each distance is calculated 1000 times. The TOA estimation performance is shown in Fig. 9, and the successful rate of first path detection is shown in Fig. 10. At last, the performance of TOA estimation under AWGN with SNR $\in [-5, 5]$dB is presented in Fig. 11.

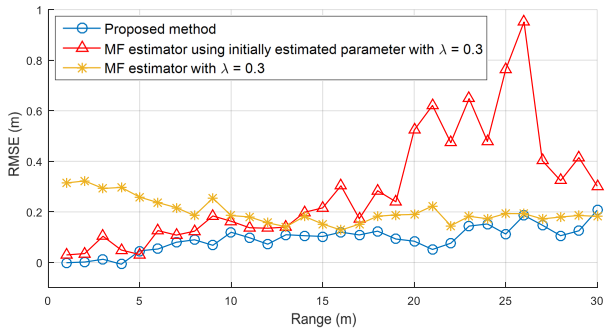


Fig. 9. RMSE performance of TOA estimation.

The result of Fig. 9 is obtained under AWGN noise and SNR $= 0$ dB. The performance of conventional MF estimator with the fixed threshold and our proposed method is compared. We can find that our proposed method can achieve a higher accurate TOA estimation and more stable than other MF estimators. For the method of MF estimator using prior knowledge, performance can be greatly improved by constructing the reference signal with initially estimated parameters, when the ranging distance is within 15 m. However, with ranging distance exceeding 20 m, TOA estimation accuracy decreases rapidly and becomes unstable. Within shorter ranging distance, the SNR of received signal is higher than the longer distance. Because the reliability of initially estimated parameter is decreased as the increase of ranging distance, the performance of initial parameter estimation based on ML estimator decreases rapidly when ranging distance increases.

The performance of the MF estimator using prior knowledge of the received signal seems abnormal. This phenomenon is caused by signal parameter setting in this simulation. We can check the TOA estimation result in Fig. 8. Due to prior knowledge of initial frequency and modulation frequency being smaller than actual value, that is $f_0 < f_0'$ and $k_0 < k_0'$, time delay estimation $\hat{\tau}_0 = 476\Delta$ is smaller than the true-value. While the ranging distance is larger and SNR is lower, the energy of earlier arrived component will be greatly impaired which makes them unable to be detected in the MF output. The TOA estimation result will become larger and larger, and gradually approach the true value, and exceed it finally. Under this condition, the RMSE of the TOA estimation will appear as a decreased trend, which is shown in Fig. 9, and an increased trend soon afterward.

Our proposed method can achieve a higher TOA estimation performance and stability. This is benefiting from the successful rate of the first path detection. The performance of detecting the first path component is also investigated and shown in Fig. 10.

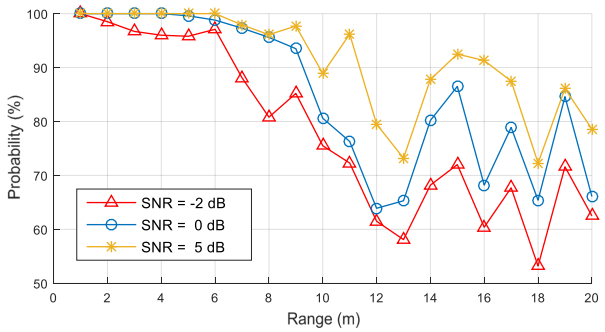


Fig. 10. Successful rate of first path detection under different SNR.

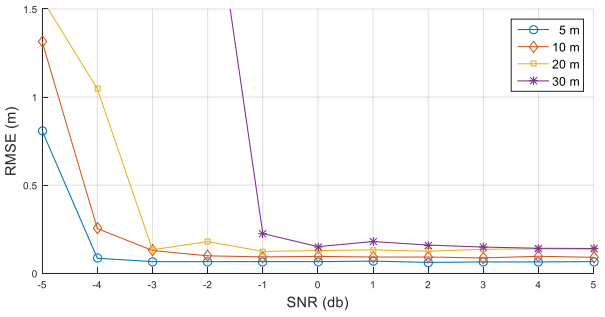


Fig. 11. Performance of TOA estimation under different SNR.

The successful rate of first path detection is investigated at three often encountered noise levels in practical applications, which are -2 dB, 0 dB and 5 dB. We can find that the performance of first path detection is satisfactory within 10 m ranging distance. The successful detection rate is higher than 75% even under SNR $= -2$ dB. While ranging distance is longer than 10 m, detection performance starts decreasing as the raging distance increases. The maximum ranging area shown

in Fig. 10 is 20 m . When the ranging distance reaches 30 m, the first path is very hard to be detected, where probability will become lower than 10%. This is because the first path is submerged in noise and can't be audible due to the lower SNR.

The TOA estimation performance under different SNRs is shown in Fig. 11. Four distances are chosen to evaluate TOA performance under AWGN with SNR $\in [-5, 5]$dB. We can achieve an accuracy of ranging error smaller than 30 cm within 30 m when SNR ≥ -1 dB, and within 20 m when SNR ≥ -3 dB. Furthermore, the time consumption of TOA estimation is from 65 ms to 95 ms, which is counted by the *tic* and *toc* function of Matlab on a computer with 3.2 GHz 4-core processor and 12 G RAM. Due to the SNR of the received signal being usually larger than -2 dB in practical applications, our proposed method can meet the demand of accurate TOA estimation of chirp signal for low-cost acoustic ranging in a real world.

V. Experiment and Results

Experiments are conducted in an underground car park near Yuquan Campus of Zhejiang University to test the performance of the proposed method in the real world. The size of this car park is $86 \times 18 \times 3.5$ (m). The scenario and device used in this experiment are shown in Fig. 12. The device is a specially designed low-cost module with acoustic broadcasting and sampling function. The main chips of acoustic part are STM32F407 and WM8978, which are very cheap and easily accessible. Microphones and speakers also use low-cost MEMS component. Total cost of them is less than 5$. To realize one-way ranging, the local time of transmitter and receiver is synchronized by ZigBee module which can provide a high time synchronization precision for low propagation speed of acoustic signal. Two devices are used during this test, where one of them is used as the transmitter and another one is used as the receiver.

(a) Underground car park (b) Device

Fig. 12. Test scenario and device used in experiment.

The parameters of acoustic chirp signal used in this test is the same as source signal, which is listed in Table I. The tested distances are uniformly distributed from 1 m to 30 m. Sound pressure level (SPL) of background noise measured in the underground car park is 40 dB and the SPL of source signal is measured as 65 dB. RMSE of TOA estimation is calculated with 100 signal samples at each selected position.

The optimal value of λ is determined by searching the parameter space between 0 and 1 with interval 0.1. Test results

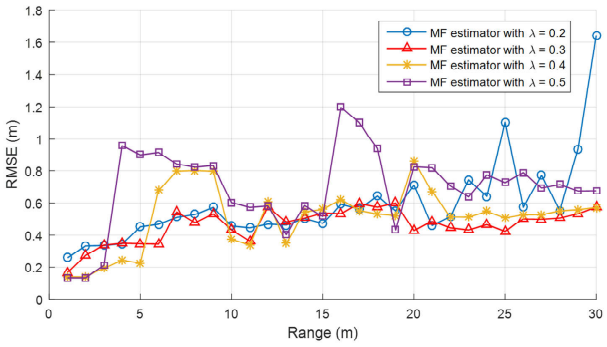


Fig. 13. Performance of MF estimator with different scaling factors.

under $\lambda = \{0.2, 0.3, 0.4, 0.5\}$ are shown in Fig. 13. When λ equals to 0.1 or is larger than 0.5, the performance decreases rapidly. The performance of $\lambda = 0.2$ and 0.3 are close with each other within short ranging distance. At some distance, the performance of $\lambda = 0.2$ are better than $\lambda = 0.3$. But when the ranging distance is longer than 21 m, the performance of $\lambda = 0.3$ are better than $\lambda = 0.2$. Comprehensively considering the TOA estimation precision and stability, the performance of MF estimator is better when the value of λ is set as 0.3.

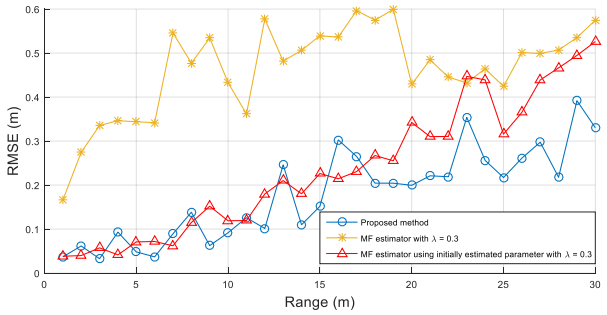


Fig. 14. Performance of TOA estimation in the real world.

From test result which is shown in Fig. 14, we can find that the proposed method could provide a higher TOA estimation precision than the conventional threshold-based MF estimator using prior knowledge and initially estimated parameter, where the RMSE of TOA estimation is smaller than 33.92 cm within 30 m ranging distance. The performance of the proposed method and MF estimator using initially estimated parameters are close to each other within 10 m, with an accuracy smaller than 12 cm, because the initially estimated parameter is still reliable for the MF estimator. While the ranging distance increases, the performance of proposed method gradually exceeds MF estimator with initially estimated parameter. Thus, through testing in the real world, we can conclude that proposed method could realize an accurate and robust TOA estimation for acoustic ranging in dense multipath environments.

In addition, the performance of our proposed method was tested and verified in the new 'Microsoft indoor localization competition' of IPSN, CPS Week 2018, Porto. The evaluated prototype systems, 'RA^{2}Loc' [35] and 'AidLoc' [36], which are designed based on a reduced version of the proposed method, achieved an average 3 D localization error of 70 cm and 71 cm,

respectively, under a dynamic real scenario [37].

VI. Conclusion

In this paper, a novel TOA estimation method is proposed for acoustic chirp signal based on an iterative 'cleaning' process in FrFD, aiming to realize robust low-cost acoustic ranging in dense multipath environments. Firstly, it has been confirmed that the conventional TOA estimator based on MF method usually can't offer sufficient TOA estimation accuracy in situations of dense multipath environments, especially for low-cost acoustic ranging. Threshold-based MF estimator can't obtain a reliable TOA estimation. Secondly, the frequency shift of transmitted signal via low-cost infrastructure widely exists due to the sampling frequency offset between the transmitter and the receiver. It makes the prior knowledge of source signal no longer reliable for the MF estimator. To mitigate these limitations, this novel TOA estimation method is introduced in FrFD.

The core of this method is an iterative 'cleaning' process based on FrFT to detect and extract the first path component from the received signal. To further improve the stability of the first path detection, the slack threshold and strict threshold are used during the iterative process. The design approach of window functions for band-pass filter and band-elimination filter in FrFD are also introduced in detail. Meanwhile, the optimal parameters of window function are given. In order to make proposed method used in practical applications, six simple and easy calculated termination criteria are proposed to monitor the iterative process. At each iterative procedure, the parameters of input signal are firstly estimated to guarantee that all the operations are within the optimal FrFD. The current strongest multipath component is eliminated under the monitoring of termination strategy. When the iterative 'cleaning' process is done, the outputs include the enhanced first path component and its estimated parameters. Using the estimated parameter of extracted first path component, we can construct an optimal reference signal for the MF estimator, and a more accurate TOA estimation of acoustic chirp signal can be conveniently obtained in dense multipath environments for low-cost acoustic ranging. The results obtained from numerical simulations and experimental investigations verified that, for acoustic chirp signal TOA estimation, the accuracy of the proposed method is superior to those obtained by the conventional MF estimator.

References

[1] A. Ens, F. Höflinger, J. Wendeberg, J. Hoppe, R. Zhang, A. Bannoura, L.M. Reindl, C. Schindelhauer. "Acoustic Self-Calibrating System for Indoor Smartphone Tracking," *Int. J. Navig. Observ.*, pp. 1–15, 2015.

[2] X. Wang, C. Zhang, F. Liu, Y. Dong, X. Xu. "Exponentially Weighted Particle Filter for Simultaneous Localization and Mapping Based on Magnetic Field Measurements," *IEEE Trans. Instrum. Meas.*, vol. 66, pp. 1658–1667, 2017.

[3] H. Liu, H. Darabi, P. Banerjee, J. Liu. "Survey of Wireless Indoor Positioning Techniques and Systems," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 37, pp. 1067–1080, 2007.

[4] A.D. Cheok, L. Yue. "A Novel Light-Sensor-Based Information Transmission System for Indoor Positioning and Navigation," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 1, pp. 290-299, 2010.

[5] A.D. Angelis, A. Moschitta, P. Carbone, M. Calderini, S. Neri, R. Borgna, M. Peppucci. "Design and Characterization of a Portable Ultrasonic Indoor 3-D Positioning System," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 10, pp. 2616-2625, 2015.

[6] L. Zhang, D. Huang, X. Wang, C. Schindelhauer, Z. Wang. "Acoustic N-LOS Identification Using Acoustic Channel Characteristics for Smartphone Indoor Localization," *Sensors.*, vol. 17, pp. 1–22, 2017.

[7] C. Peng, G. Shen, Y. Zhang. "BeepBeep:A High-Accuracy Acoustic-based System for Ranging and Localization Using COTS Devices," *ACM Trans. Embedded Comput. Syst.*, vol. 11, no. 4, pp. 1-29, 2012.

[8] C. Tan, X. Zhu, Y. Su, Y. Wang, Z. Wu, D. Gu. "A low-cost centimeter-level acoustic localization system without time synchronization," *Measurement*, vol. 78, pp. 73–82, 2016.

[9] K. Liu, X. Liu, X. Li. "Guoguo: Enabling Fine-Grained Smartphone Localization via Acoustic Anchors," *IEEE Trans. Mob. Comput.*, vol. 15, pp. 1144–1156, 2016.

[10] S.I. Lopes, J.M.N. Vieira, J. Reis, D. Albuquerque, N.B. Carvalho. "Accurate smartphone indoor positioning using a WSN infrastructure and non-invasive audio for TDoA estimation," *Pervasive Mob. Comput.*, vol. 20, pp. 29–46, 2015.

[11] D. Dardari, C.C. Chong, M.Z. Win. "Threshold-Based Time-of-Arrival Estimators in UWB Dense Multipath Channels," *IEEE Transactions on Communications*, vol. 56, pp. 1366–1378, 2008.

[12] Mark A. Richards. Fundamentals of Radar Signal Processing, *McGraw-Hill Professional*, USA, pp. 197, 2005.

[13] L. Stoica, A. Rabbachin, I. Oppermann. "A low-complexity noncoherent IR-UWB transceiver architecture with TOA estimation," *IEEE Transactions on Microwave Theory & Techniques*, vol. 54, pp. 1637–1646, 2006.

[14] C.W. Chung, H. Dong. "An accurate ultra wideband (UWB) ranging for precision asset location," *Ultra Wideband Systems and Technologies, 2003 IEEE Conference on.*, Reston, USA, Nov., 2003, pp. 389-393.

[15] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, M.Z. Win. "Ranging With Ultrawide Bandwidth Signals in Multipath Environments," *Proceedings of the IEEE*, vol. 97, pp. 404–426, 2009.

[16] I. Guvenc, Z. Sahinoglu. "Threshold selection for UWB TOA estimation based on kurtosis analysis," *Communications Letters IEEE*, vol. 9, pp. 1025–1027, 2005.

[17] A.Y.Z. Xu, E.K.S. Au, A.K.S. Wong, Q. Wang. "A Novel Threshold-Based Coherent TOA Estimation for IR-UWB Systems," *IEEE Transactions on Vehicular Technology*, vol. 58, pp. 4675–4681, 2009.

[18] J.Y. Lee, S. Yoo. "Large error performance of UWB ranging in multipath and multiuser environments," *IEEE Transactions on Microwave Theory & Techniques*, vol. 54, pp. 1887–1895, 2006.

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

[20] H. Farrokhi. "TOA Estimation using Music Super-Resolution Techniques for an Indoor Audible Chirp Ranging System," *IEEE International Conference on Signal Processing and Communications*, Dubai, UAE, Nov., 2008, pp. 987-990.

[21] T. Okano, S. Kidera, T. Kirimoto. "Super Resolution TOA Estimation Algorithm with Maximum Likelihood ICA Based Pre-Processing," *IEICE Transactions on Communications, vol. E96.B, issue 5, pp.1194-1201, 2013.*

[22] I. Sharp, K. Yu. "Indoor TOA Error Measurement, Modeling, and Analysis," *IEEE Trans. Instrum. Meas.*, vol. 63, no. 9, pp. 2129-2144, 2014.

[23] H. Kuttruff. Room Acoustic, 4th ed., *Spon Press*, London, UK, pp. 89–114, 2006.

[24] J. Chen, J. Benesty, Y. Huang. "Time delay estimation in room acoustic environments: An overview," *EURASIP J. Appl. Signal Process.*, Dec., 2006, 2006:026503.

[25] K.K. Sharma, S.D. Joshi. "Time delay estimation using fractional Fourier transform," *Signal Processing*, vol. 87, no.5, pp. 853-865, 2007.

[26] T. Zhou, H. Li, J. Zhu, C. xu. "Subsample time delay estimation of chirp signals using FrFT," *Signal Processing*, vol. 96, no. 5, pp. 110-117, 2014.

[27] T.J. Abatzoglou. "Fast maximum likelihood joint estimation of frequency and frequency rate," *IEEE Trans. on AES*, vol. 22, no. 6, pp. 708-715, 1986.

[28] S. Peleg, B. Porat. "Estimation and Classificatoin of Polynomial-phase Signal," *IEEE Trans. on Information Theory*, vol. 37, no. 2, pp. 423-430, 1991.

[29] M.S. Wang, A.K. Chan, C.K. Chui. "Linear Frequency Modulated Signal Detection Using Radon-Ambiguity Transform," *IEEE Trans. on Signal Processing*, 1998, vol. 46, no. 3, pp. 571-586, 1998.

[30] P.M. Baggenstoss. "Recursive Decimation/Interpolation for ML Chirp Parameter Estimation," *IEEE Transactions on Aerospace & Electronic Systems*, vol. 50, no. 50, pp. 445-455, 2014.

[31] K.B. Jennison. "Detection of Polyphase Compression Waveforms Using the Radon-Ambiguity Transform," *IEEE Trans. on AES*, vol. 39, no. 1, pp. 335-343, 2003.

[32] L.B. Almeida. "The fractional Fourier transform and time-frequency representations," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 3084-3091, 1994.
[33] H.M. Ozaktas, O. Arikan, M.A. Kutay, G. Bozdagt. "Digital computation of the fractional Fourier transform," *IEEE Transactions on Signal Processing*, vol.44, no. 9, pp. 2141-2150, 1996.
[34] E.A.P. Habets. "Room Impulse Response Generator," *chnische Universiteitndhoven*, vol. 25, no. 2, pp. 1-17, 2010.
[35] L. Zhang, M. Chen, X. Wang, Z. Wang. "RA^{2}Loc: A Robust Accurate Acoustic Indoor Localization System," *Microsoft Indoor Localizatoin Competition*, Porto, Portugal, Apr., 2018, pp. 1-2.
[36] M. Chen, L. Zhang, X. Wang, Z. Wang. "AidLoc: An Accurate Acoustic 3D Indoor Localization System," *Microsoft Indoor Localizatoin Competition*, Porto, Portugal, Apr., 2018, pp. 1-2.
[37] Available online:https://www.microsoft.com/en-us/research/event/microsoft-indoor-localization-competition-ipsn-2018/,accessed on 12 April 2018.

Lei Zhang received his B.Eng. and M.Sc. degrees in mechatronic engineering from Chang'an University, Xi'an, China, in 2007 and 2013, respectively. He is currently pursuing a Ph.D. degree in control science and engineering at Zhejiang University, Hangzhou, China. Lei's academic interests include machine learning, acoustic signal processing, time-frequency analysis, and indoor positioning and tracking.

Minlin Chen received his B.Eng. degrees in mechanic engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2018. He is currently pursuing a Master degree in control engineering at Zhejiang University, Hangzhou, China. Minlin's academic interests include machine learning, acoustic signal processing, and indoor positioning and tracking.

Xinheng Wang (M'04-SM'14) received the B.E. and M.Sc. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1991 and 1994, respectively, and the Ph.D. degree in electronics and computer engineering from Brunel University London, Uxbridge, U.K., in 2001.

He is currently a Professor of Computing with the School of Computing and Engineering, University of West London, London, U.K. He is the investigator/co-investigator of more than 20 research projects sponsored from EU, UK EPSRC, Innovate UK, China NNSFC, and industry. He holds seven patents and has authored or co-authored over 150 referred papers. His research has led to a few commercial products in condition monitoring, wireless mesh networks (www.swanmesh.com), and healthcare. His current research interests include indoor positioning, Internet-of-Things, and big data analytics for smart airport services, where he has developed the world's first smart trolley with an industry partner.

Zhi Wang received the B.E. degree in mechanism from Shenyang Jianzhu University, Liaoning, China, in 1991, the M.Sc. degree in mechanical engineering, Southeast University, Nanjing, Jiangsu, China, in 1997, and the Ph.D. degree in automated control, Shenyang Institute of Automation, China Academy of Science, Liaoning, China, in 2000.

He is currently an associate professor at State Key Laboratory of Industrial Control, Zhejiang University, China. His main research focus is on acoustic signal and array processing, sparsity signal and compressive sensing, localization and tracking of mobile target, crowdsourcing, mobile computing and industrial IoT protocol.

Prof. Zhi Wang is the principal investigator or co-investigator for more than 20 projects from Zhejiang province, China government, international organizations, and industry. Prof. Zhi Wang has co-authored over 100 publications in international journals (such as IEEE TSP, IEEE TPDS, IEEE TMC, IEEE TII) and conferences such as IEEE INFORCOM, IEEE SECON, IEEE DCOSS, IEEE WFCS). He has served as advising board and editor of IJDSN, as General Co-Chair (IEEE DCOSS 2012) and as TPC member (IEEE WFCS, IEEE ETFA, Sensornets et al.) for a number of cyber physical systems and sensor networks related international conferences. Prof. Wang is the committee member for China Computer Federation Sensor Network Technical Committee and China National Technical Committee of Sensor Network Standardization, and is a member of IEEE and ACM.