MIMO Radar for Advanced Driver-Assistance Systems and Autonomous Driving: Advantages and Challenges

Article in IEEE Signal Processing Magazine - July 2020
DOI: 10.1109/MSP.2020.2978507

3 authors, including:

Shunqiao Sun
University of Alabama
19 PUBLICATIONS  210 CITATIONS

All content following this page was uploaded by Shunqiao Sun on 03 July 2020.
The user has requested enhancement of the downloaded file.
Important requirements for automotive radar are high resolution, low hardware cost, and small size. Multiple-input, multiple-output (MIMO) radar technology has been receiving considerable attention from automotive radar manufacturers because it can achieve a high angular resolution with relatively small numbers of antennas. For that ability, it has been exploited in the current-generation automotive radar for advanced driver-assistance systems (ADAS) as well as in next-generation high-resolution imaging radar for autonomous driving. This article reviews MIMO radar basics, highlighting the features that make this technology a good fit for automotive radar and reviewing
important theoretical results for increasing the angular resolution. The article also describes challenges arising during the application of existing MIMO radar theory to automotive radar that provide interesting problems for signal processing researchers.

Introduction

Radar technology has been used in defense, civilian, and commercial applications since World War II. Surveillance radars have been in service for a number of decades. Ground-based air surveillance radars, such as air traffic control radar, are usually pulse radar systems with maximum detectable ranges of more than 100 km [1]. Airborne surveillance radars carried by aircraft and unmanned aerial systems, such as the joint surveillance target attack radar system, can run in ground moving target indication and synthetic aperture radar imaging modes [2] to detect moving and stationary targets on the ground at more than 250 km, respectively. To achieve a satisfactory return-signal power, the transmit power of such radar can be several kilowatts. Large phased-array antennas have been deployed to achieve electronic beam scanning. Since the late 1990s, radar sensors have found widespread applications in ADASs, such as adaptive cruise control (ACC) and automatic emergency braking (AEB) [3], [4]. More recently, radar has emerged as one of the key technologies in autonomous driving systems, providing environmental perception in all weather conditions [5], [6]. Some of today’s self-driving cars, such as Zoox, have more than 10 radars, providing 360° surround sensing (see Figure 1). Differing from ground-based and airborne surveillance radars, automotive radars have a small size (multi-inch by multi-inch), short range (within multihundreds of meters), low power consumption (multiwatt), and low cost. They are integrated behind the vehicle bumper or windshield, operating in a highly dynamic environment with multipath, as shown in Figure 2.

Figure 3 presents a typical radar perception configuration for autonomous driving. Four short-range radar (SRR) sensors with an azimuth FoV of $[-75°, 75°]$ and a detection range of 45 m are deployed at the four corners of the vehicle for blind-spot detection, cross traffic alerts, and so forth. Two midrange radar (MRR) sensors with an azimuth FoV of $[-40°, 40°]$ and a detection range of 100 m are deployed at both the front and rear sides for lane-change assistance and AEB. A long-range radar (LRR) sensor with an azimuth FoV of $[-15°, 15°]$ and a detection range of 250 m is deployed at the front for the ACC. All radar sensors in a vehicle are connected to an electronic control unit for further processing, such as radar tracking and sensor fusion [7].
For control critical functions, such as AEB, distant obstacles need to be detectable at a high angular resolution. Therefore, automotive radar for ADAS and autonomous driving needs to have a high angular discrimination capability. Employing a large antenna array would improve the angular resolution; however, the resulting large package size would make integration on the vehicle difficult. While for conventional phased-array radar, a small package size implies a low angular resolution, for MIMO radar [8]–[10], the package size is not a limiting factor. This is because MIMO radar can synthesize virtual arrays with a large aperture using only a small number of transmit and receive antennas.

This advantage has been exploited by almost all major automotive suppliers in their different types of radar products, such as SRR, MRR, and LRR [11]–[14]. For Level 4 and Level 5 (L4/L5) autonomous driving, both the azimuth and elevation angular resolution of automotive radar need to be less than 1°. High-resolution imaging radar with hundreds of virtual array elements is currently being developed to produce the so-called point clouds, which are groups of points detected by radar that represent the object’s shape [11], [15], [16]. A variety of signal processing tools, e.g., the fast Fourier transform (FFT), short-time Fourier transform, filtering, and beamforming, have been adopted in automotive radar to obtain target features, such as a micro-Doppler spectrum of pedestrians [17]–[20] and a range-Doppler spectrum of the surrounding environment. Machine learning algorithms and deep neural networks [21] have also been applied in automotive radar for target recognition and classification.

In summary, to meet the requirement for ADAS and especially L4/L5 autonomous driving, automotive radar needs to have a high angular resolution, small package size, and low cost. Since it is meant to operate in highly dynamic environments, it needs to detect obstacles fast, especially for time critical functions, such as AEB. In this article, we discuss MIMO radar with millimeter-wave (mm-wave) technology as a means of achieving the aforementioned requirements. We review key issues in state-of-the-art frequency-modulated continuous-waveform (FMCW) MIMO radar, different strategies for achieving waveform orthogonality, virtual array synthesis, and high-resolution angle finding methods for both uniform linear arrays and sparse linear arrays. We discuss MIMO radar in the light of high-resolution imaging radar for L4/L5 autonomous driving. We also highlight challenges in designing automotive MIMO radar, such as angle finding in the presence of multipath, waveform-orthogonality strategies in FMCW radar and pulse-modulated CW (PMCW) radar, and the mitigation of radar mutual interference. We hope this article serves as a tutorial on automotive MIMO radar and provides interesting research problems to signal processing researchers.

**State-of-the-art automotive FMCW radar**

Mm-wave technology has found great applicability in automotive radar. The typical frequency band of mm-wave automotive radar is 76–81 GHz. The high frequencies facilitate small enough antennas that can fit behind the bumper of the vehicle. Also, the wide available bandwidth enables a high target-range resolution [22]. State-of-the-art automotive radar transmits FMCW at mm-wave frequencies, which supports high-resolution target-range and velocity estimation at a much lower cost than lidar technology. Automotive MIMO radar uses FMCW waveforms along with some mechanism that guarantees waveform orthogonality. This section introduces the principles of FMCW. Strategies for achieving waveform orthogonality based on FMCW waveforms are discussed in the “Introduction of Automotive Radar With MIMO Radar Technology” section.

An FMCW waveform, also referred to as a chirp, is a complex sinusoid whose frequency increases linearly with time \( t \in [0, T] \), i.e., \( f_r(t) = f_c + (B/T) t \), where \( B \) is the signal bandwidth and \( f_c \) is the carrier frequency. FMCW radar transmit chirps in a periodic fashion, with a period referred to as the pulse repetition interval (PRI). The frequency of an FMCW signal through multiple periods, with a PRI equal to \( T \), is displayed in Figure 4. The target echo at the radar receiver contains a delayed and attenuated copy of the transmitted chirp. For a target at range \( R \), moving with a radial speed of \( v \), the delay equals \( \tau = (2(R + vt))/c \), where time \( t \) spans multiple periods and \( c \) is the speed of light. The received signal is mixed with the transmitted chirp, which results in a complex sinusoid known as the beat signal. The beat-signal frequency equals \( f_b = f_c + f_d \), where \( f_b = 2RB/(Tc) \) is the range frequency and \( f_d = (2v/c)f_c \) is the Doppler frequency.

![FIGURE 4. The FMCW radar chirps. The range and Doppler estimations are performed using a 2D FFT.](image-url)
The process of obtaining the beat signal is implemented in the radio frequency domain by a mixer, followed by a bandpass filter (BPF) with the maximum cutoff frequency \( f_{\text{B}}^{\text{max}} \); the latter filter is used to remove signals with frequencies outside the band of interest, which also places a limit on the maximum detectable range.

The estimation of the beat frequency is implemented in the digital domain, after the sampling of the beat signal. In automotive scenarios, the maximum detectable range, \( R_{\text{max}} \), is hundreds of meters. It holds that \( (2R_{\text{max}}/c) \ll T \), and thus \( f_{\text{B}} \ll B \). Since it typically holds that \( f_0 \ll f_{\text{B}} \), the beat frequency is much smaller than \( B \), and therefore a low-speed analog-to-digital converter (ADC) can be used to sample the beat signal. The time during one period or chirp is usually referred to as the *fast time*, while the time across multiple periods or chirps is referred to as the *slow time*. Thus, if we sample the beat signal and put the samples of each chirp in the columns of a matrix, the row indices of that matrix correspond to the fast time and the column indices to the slow time (see Figure 4). In automotive scenarios, \( f_0 \ll f_{\text{B}} \), therefore, \( f_0 \) can be taken as constant within each chirp. Thus, by applying FFTs on the sampled beat signal along the fast time, one can identify \( f_{\text{B}} \) based on which of the target’s ranges can be obtained as \( R = c f_{\text{B}} T/(2B) \).

To obtain the target’s Doppler frequency, a second FFT operation is subsequently carried out along the slow time (the range frequency \( f_{\text{B}} \) is the same across the slow time). The PRF is \( f_{\text{PRF}} = 1/T_{\text{PRI}} \). To avoid Doppler ambiguity, it is desired that \( f_{\text{PRF}} \geq 2f_0 \). Thus, the maximum unambiguous detectable radial speed of FMCW radar is \( v_{\text{max}} = c/(4f_{\text{PRF}}) \) [23]. The application of these two FFTs is equivalent to a 2D FFT of the beat signal in the fast and slow times, and the result is called the range–Doppler spectrum. Range and Doppler detection can be performed using conventional thresholding-based methods applied to the 2D range–Doppler spectrum, such as the constant false alarm rate detector [24] or the recently proposed deep neural network-based detector [25]. Via the 2D FFT, the targets can be separated in the range and Doppler domains. Since the number of targets within the same range–Doppler bin is small, angle finding can be carried out using sparse sensing techniques, such as compressive sensing. The details of angle finding approaches will be covered in the “Angle Finding in Automotive MIMO Radar” section.

The 2D FFT operation used for beat-frequency estimation can be computed with low-cost digital signal processors (DSPs) and field-programmable gate arrays. The range resolution depends on the beat-frequency resolution. Since the latter is estimated based on a signal of time duration approximately equal to \( T \), the beat-frequency resolution is \( 1/T \), and thus the range resolution is \( c/(2B) \). One can see how the range resolution benefits from a high bandwidth. Thus, the low hardware cost, coupled with the high range resolution that can be achieved, make FMCW radar very popular in automotive radar applications.

As an example, an LRR FMCW radar operating at carrier frequency \( f_c = 77 \text{ GHz} \) has a typical pulse duration of \( T = 50 \mu\text{s} \) and bandwidth of \( B = 150 \text{ MHz} \). Assume the maximum detectable range is 250 m. Then, the maximum range frequency is \( f_{\text{B}}^{\text{max}} = 5 \text{ MHz} \). For a moving target with a maximum speed of \( v = 120 \text{ m} / \text{s} \), the maximum Doppler frequency is \( f_{\text{D}}^{\text{max}} = 27.4 \text{ kHz} \). The beat signal can be sampled with a low-cost ADC with a sampling rate as low as \( f_s = 12 \text{ MHz} \). For a PRI equal to \( T \), and based on the preceding, the PRF is \( f_{\text{PRF}} = 1/T_{\text{PRI}} = 20 \text{ kHz} \). It can be seen that, in this example, \( f_{\text{PRF}} < 2f_{\text{B}}^{\text{max}} \), and therefore the Doppler frequency will be aliased, resulting in Doppler ambiguity. In practice, chirp sequences with different PRIs are transmitted to resolve the Doppler ambiguity.

As discussed, FMCW radar has the desirable ability to estimate the target range and Doppler at a low cost. However, FMCW radar also faces several challenges. Those include frequency sweep nonlinearity [26]–[28] arising from the phase noise of the local oscillator (LO) and imperfect phase-locked loop circuits and voltage-control oscillators. Sweep nonlinearity has a significant impact on radar range estimation [26].

A challenge common to all radars is to maintain a high dynamic range, which is the ratio between the maximum and minimum possible successfully received power. Since the path loss is inversely proportional to the fourth order of the range, targets at long distances typically yield weaker reflections compared to targets at closer distances. Therefore, a high dynamic range is required to accommodate targets at a wide range of distances. In FMCW radar, because the waveforms are transmitted continuously, the receiver always receives the signals of the transmitters via a direct path, referred to as leakage. Leakage may saturate the low-noise amplifier in the front end of the radar receiver [29]. It can also cause high peaks at the first few bins of the range FFT, thus reducing the radar’s dynamic range.

In many pulse radars, sensitivity time control has been widely used to control the gain of targets at near and far distances by exploiting the relationship between the time delay and target range. As discussed, the range frequency of the beat signal is proportional to the target range. Therefore, in FMCW radar, an analog gain-control logic can be implemented in the frequency domain to reduce the receiver gain of close-range targets while continually increasing the gain as the target’s range increases [30]. Automatic gain-control logic can be implemented in a BPF so that the receiver gain increases within the detectable range [31]. Leakage between the transmit and receive antennas can be suppressed via sufficient isolation between the transmit and receive antennas through antenna design in the analog domain [30], [32]. Further, the transmit antenna leakage effect in the mixer output can be suppressed via a BPF, with the lowest cutoff frequency corresponding to a range of tens of centimeters to 1 m. The BPF also helps suppress the multibounce signals between automotive targets, the radome, and the vehicle bumper. Finally, in the

Radar technology has been used in defense, civilian, and commercial applications since World War II.
complete receiving chain, an $N$-bit ADC can provide an additional 6$N$-dB dynamic range.

In the signal processing chain of FMCW radar, a range FFT length of $N_R$ and a Doppler FFT length of $N_D$ can provide a signal-to-noise ratio (SNR) improvement of $10 \log_{10}(N_R N_D)$ dB [33]. This SNR improvement is considered as a processing gain, which significantly benefits angle finding that will be discussed in the “Angle Finding in Automotive MIMO Radar” section.

**Introduction of automotive radar with MIMO radar technology**

In state-of-the-art automotive FMCW radar, the range and Doppler parameters of targets can be estimated by using a single receive antenna. However, to estimate the angle parameter of targets, a receive antenna array is needed. In MIMO radar, the transmit antennas transmit FMCW sequences in a way that guarantees their orthogonality. At each receive antenna, the contribution of each transmit antenna is extracted by exploiting waveform orthogonality. For $M_t$ transmit antennas and $M_r$ receive antennas, a virtual array with $M_t M_r$ elements can be synthesized. The array response of the synthesized array, i.e., the array corresponding to a MIMO radar, can be expressed as $\mathbf{a}(\theta_t) \otimes \mathbf{a}(\theta_r)$, where $\mathbf{a}(\theta_t)$ and $\mathbf{a}(\theta_r)$ are the transmit- and receive-array steering vectors corresponding to a direction of departure (DoD) $\theta_t$ and direction of arrival (DoA) $\theta_r$, respectively. Here, $\otimes$ denotes the Kronecker product.

Figure 5 shows two MIMO radar configurations with $M_t = 2$ transmit and $M_r$ receive antennas [23]. One is in interleaved mode with $d_i = 2 d_t$ and $d_l = \lambda / 2$, and the other is in stacked mode with $d_i = M_t d_t$ and $d_l = \lambda / 2$. Here, $\lambda = c / f_0$ is the wavelength. It can be verified that for both array configurations, the synthetic virtual arrays are equivalent to uniform linear arrays (ULAs) with $M_t M_r$ elements and a spacing of $d_i$ when $\theta_t = \theta_r$.

Because MIMO radar transmits orthogonal waveforms, when isotropic array elements are used, the array beampattern (also referred to as the MIMO radar array factor) is omnidirectional. Thus, MIMO radar loses the coherent array-processing-gain advantage enjoyed by traditional phased-array radar systems [10]: the SNR of the array response at a given angular direction is less than that of phased-array radar with transmit beamforming. Still, in the automotive application scenario, the high-resolution angle finding ability of MIMO radar coupled with its low cost are viewed as more important than the loss of coherent processing gain.

Virtual array synthesis in automotive FMCW radar using MIMO radar technology relies on the separability of the transmit signals of the different antennas. The separation is easier when the transmit signals of different antennas are orthogonal. In the following, we review techniques to achieve waveform orthogonality while transmitting FMCW, such as TDM, DDM, and frequency-division multiplexing (FDM).

**Waveform orthogonality via TDM**

In TDM MIMO radar [23], [35]–[38], only one transmit antenna is scheduled to transmit at each time slot. In Figure 6, a signal processing example of a MIMO radar in TDM is given by NXP Semiconductors [34], where $M_t = 2$ transmit antennas emit FMCW chirps alternatively. The switch delay between transmit antennas is $\Delta T = T_{PRI}$. At each receive antenna, range FFTs of length $N_r$ are conducted for each chirp, and the FFT outputs of $2 N_d$ chirps are assembled in two matrices corresponding to odd and even chirp sequences, respectively. The receive array corresponding to the odd and even chirp sequences form two subarrays, which can be used to synthesize a virtual array according to interleaved or stacked configurations.

For a moving target with a velocity of $v$, the switching delays of the transmit antennas introduce a target phase migration from chirp to chirp, which is defined as $\phi = 2 \pi f_0 \Delta T = (4 \pi / \lambda) v \Delta T$. As a result, the virtual array pattern would be distorted [23]. In Figure 6, the phase difference between corresponding columns in the two matrices is $\phi = (4 \pi / \lambda) \nu T_{PRI}$. If $\nu = -0.5 \nu_{\text{max}}$ and $\nu = -\nu_{\text{max}}$, where $\nu_{\text{max}}$ is the maximum unambiguous detectable radial speed and $\nu_{\text{max}} = c / (4 f_0 T_{PRI})$ (see the “State-of-the-Art Automotive FMCW Radar” section), the phase shifts are $\phi = -\pi / 2$ and $\phi = -\pi$, respectively. The array beampattern distortion is demonstrated in Figure 7 for a moving target.

![Figure 5](image-url)  
**Figure 5.** The different MIMO radar virtual array configurations [23] using the time-division multiplexing (TD) or Doppler-division multiplexing (DDM) scheme with $M_t = 2$ transmit antennas and $M_r$ receive antennas: (a) interleaved, with $d_i = 2 d_t$ and $d_l = \lambda / 2$; (b) stacked, with $d_i = M_t d_t$ and $d_l = \lambda / 2$. Different colors indicate that the transmit antenna transmits different time slots or codes.
**FIGURE 6.** An example of radar signal processing with the TDM scheme [34], where $M_t = 2$ transmit antennas alternately transmit FMCW chirp sequences. The red and green colors denote the odd and even echo chirp sequences, respectively. The range FFTs are conducted for each chirp, and the FFT outputs are stored in two matrices corresponding to odd and even sequences, respectively, for further processing.

**FIGURE 7.** The examples of the MIMO radar range and azimuth images using TDM [23], with $M_t = 2$, $M_r = 8$ for a single target with a range of 35 m and azimuth angle of $\theta = 0^\circ$. Two MIMO array configurations, interleaved (left column) and stacked (right column), are considered. (a) and (b) The radial velocity of $v = -v_{\text{max}}$. (c) and (d) The radial velocity of $v = -0.5 v_{\text{max}}$. (e) and (f) The radial velocity of $v = 0$. 

Authorized licensed use limited to: UNIV OF ALABAMA-TUSCALOOSA. Downloaded on July 03,2020 at 02:30:55 UTC from IEEE Xplore. Restrictions apply.
with a range of 35 m and azimuth angle of $\theta = 0^\circ$ [23]. Here, MIMO radar with $M_t = 2$, $M_r = 8$ operates in TDM fashion. As stated in [23] and seen in Figure 7, for the inter-leaved MIMO array configuration, when the target velocity increases, the grating lobes at the edge of the FoV show up, while the peak at the target direction decreases and totally disappears when $v = -v_{\text{max}}$. For the stacked MIMO array configuration, as the target velocity increases, the peak is slightly off the boresight, with a mirror grating lobe at the opposite direction.

The phase migration introduced by every moving target in the virtual array response needs to be compensated for before angle finding. The phase-shift estimate $\phi$ can be obtained after each target velocity has been estimated based on the 2D FFT of a single receive antenna or the noncoherent 2D FFT integration of the same subarray. For instance, in the example of Figure 6, the phase in the beam vector of the subarray obtained from the even chirps needs to be compensated for by multiplying it with $e^{-j\phi}$, while the phase in the beam vector of the subarray corresponding to the odd chirps is kept unchanged. It should be noted that in TDM MIMO radar, the pulse-repetition interval is enlarged by the transmit antenna number $M_t$. As a result, the maximum unambiguous detectable velocity, which was defined in the “State-of-the-Art Automotive FMCW Radar” section, will be reduced by a factor of $M_t$ [39].

**Waveform orthogonality via DDM**

In one look, a total of $N$ chirps (i.e., pulses) are transmitted sequentially with pulse-repetition interval $T_{\text{PR}}$. All transmit antennas simultaneously transmit the same FMCW waveform after multiplying it with a phase code that is different for each antenna and changes between pulses, i.e., $x_m(n) = e^{j2\pi f_{\text{PR}}nt}$, $m = 1, \ldots, M_t$, $n = 1, \ldots, N$ [40]. To separate the $n$th transmit signal at the $m$th receiver, after the range FFT, a slow-time Doppler demodulation is applied to all range bins corresponding to the same chirp. The Doppler demodulated outputs of $N$ chirps are assembled into a vector $s_n$. Then, the Doppler FFT is applied on the vector $s_n$. To separate the transmit signals in the Doppler domain, one of the two methods described next can be applied.

The first approach is to design phase codes such that the Doppler FFT of the interference $e^{j2\pi (a_n - a_m)}$ is shifted to a frequency that is higher than the maximum detectable Doppler frequency $f_{\text{D}}^{\text{max}}$. Therefore, a low-pass filter (LPF) can be applied to remove the interference [40]. One example of such phase codes is

$$a_m(n) = a_m n, \quad m = 1, \ldots, M_t, \quad n = 1, \ldots, N, \quad (1)$$

where the starting phase $a_m$ is linear across different transmit antennas, i.e., $a_m = a_m n$. Figure 8(a) shows the range and Doppler spectra of a target with a range of 75 m and velocity of 10 m/s. Automotive MIMO radar has two transmit antennas and the transmit phase codes given in (1), with $a_0 = 1$ and $N = 512$. It can be seen that signals from different transmit antennas are shifted to a higher Doppler spectrum, which can be removed via an LPF in the Doppler domain. With this approach, the radar pulse repetition frequency $f_{\text{PR}}$ should be larger than $M_t f_{\text{D}}^{\text{max}}$ [41]. Thus, if the $f_{\text{PR}}$ is kept unchanged, the maximum detectable unambiguous Doppler frequency is reduced by a factor of $M_t$. In practice, a Doppler unfolding, or de-aliasing, algorithm needs to be developed with different $f_{\text{PR}}$ in different looks.

The second approach is to design phase codes so that the Doppler FFT of the interference can be distributed into the entire Doppler spectrum as pseudo noise. It is desired to minimize the peak interference residual (PIR) in the Doppler spectrum [42], calculated using the discrete time Fourier transform for $m = 1, \ldots, M_t$, i.e.,

![FIGURE 8](image-url) The range and Doppler spectra of a target with a range of 75 m and velocity of 10 m/s. The automotive MIMO radar has two transmit antennas, and a slow time phase coding of length $N = 512$ is applied for DDM. (a) The phase shift codes defined in (1) for the range-Doppler spectrum. (b) The two Chu sequences for the range-Doppler spectrum.
where \( f \in [-(1/2)f_{PRF}, (1/2)f_{PRF}] \). Following (2), the cross correlation of the spectra of two codes needs to be flat [42] since the Fourier transform of multiplication of two codes in the time domain is equivalent to the convolution of the spectrum of one code with the time reversed and the complex conjugate of the other. The maximum autocorrelation value of an unimodular sequence of length \( N \) is \( N \). The ideal cross correlation of two unimodular sequences of length \( N \) has a magnitude of \( \sqrt{N} \). Thus, in the ideal case, according to [42], the maximum power gain of the currently transmitted signal over the other signals is \( \sqrt{N} \). For example, the maximum achievable waveform attenuation is roughly 27.1 dB for a unimodular sequence set with \( N = 512 \).

Constant-amplitude zero-autocorrelation codes are good candidates for DDM. The discrete Fourier transform of a constant-amplitude zero autocorrelation code also has a constant amplitude and zero autocorrelation [43]. One such example is the Chu sequence [44], which is defined as \( x_m(n) = e^{j(\pi/2)(m+1)n}, \quad m = 1, \ldots, M_t; \quad n = 1, \ldots, N \), where \( N \) is a prime number. In practice, the Chu sequence of the prime length is first generated and then truncated into a length for an efficient FFT. For example, we generate Chu codes of prime length 521 and truncate them to length \( N = 512 \). By calculation with the FFT, the peak interference residual defined in (2) is 1.08\(\sqrt{N}\). Therefore, the waveform attenuation for a Chu sequence of length \( N = 512 \) is approximately 26.4 dB. In Figure 8(b), we show the range and Doppler spectra of a target with a range of 75 m and velocity of 10 m/s. The automotive radar has two transmit antennas, and two Chu sequences of length \( N = 512 \) are applied for the slow-time DDM. It can be seen that the waveform attenuation is roughly 26 dB. In practice, binary phase codes are used due to hardware constraints [45]. The binary phase-code sequences are obtained via an exhaustive search such that the peak interference residual in (2) is low. As the code length increases, the search time will grow exponentially.

The benefit of slow time phase coding is that the interference from other transmitters does not affect different range bins. The range resolution is determined only by the bandwidth of the FMCW chirp. Therefore, it avoids the range sidelobe issue using fast time phase coding. However, the Doppler sidelobes would be high due to the residual of the slow time phase coding. As a result, targets with a low radar cross section (RCS), e.g., pedestrians, that are close to targets that have strong reflections, e.g., trucks, might be masked by the waveform residual. In other words, the waveform residual reduces the radar dynamic range. Given the code length, the number of phase codes with good correlation properties is limited, or, equivalently, the number of antennas that can transmit simultaneously is limited.

### Waveform orthogonality via FDM

In the FDM scheme, the transmitted signals are modulated by different carrier frequencies. According to [46], the separation of multiple transmit FMCW signals is achieved by shifting the \( m \)th transmit FMCW chirp by an offset frequency \( f_{off,m} \). If the differences between all \( f_{off,m} \) are larger than twice the cutoff frequency of the antialiasing BPF \( f_{BPF}^{\max} \), which is determined by the maximum unambiguous detectable range and Doppler, the transmitted signals can be separated at the receive end. Specifically, the received signal at each receiver is first mixed with the same starting carrier frequency \( f_c \). The separation of transmit signals in the mixer output can be implemented by a frequency shift followed by an LPF with cutoff frequency \( f_{BPF}^{\max} \). Each receiver needs to carry out such a frequency shift and filtering operation \( M_t \) times. As a result, a high range resolution can be realized using a typical FMCW chirp with a large bandwidth. Meanwhile, after the FMCW demodulation, frequency shift, and filtering operation, the FDM MIMO scheme can still utilize a low sampling rate determined by the beat signal.

Let us consider the FDM scheme in the context of the example in the “State-of-the-Art Automotive FMCW Radar” section, i.e., a FMCW LRR radar with a maximum detectable range of 250 m and a maximum detectable velocity of 120 mi/h. For bandwidth \( B = 150 \text{ MHz} \) and chirp duration \( T = 50 \mu s \), the maximum beat frequency is \( f_{BPF}^{\max} = f_{BPF}^{\max} + f_{BPF}^{\max} = 5.0274 \text{ MHz} \). Therefore, the frequency shift for the \( m \)th transmit antenna in the FDM scheme can be chosen as \( f_{off,m} = 12(m - 1) \text{ MHz} \). The intermediate frequency (IF) should have a bandwidth of \( 12 M_t \text{ MHz} \) to hold the mixer output.

### Angle finding in automotive MIMO radar

In automotive MIMO radar with \( M_t \) transmit and \( M_r \) receive antennas, a virtual uniform linear array of \( M_t M_r \) elements can be synthesized with interelement spacing \( d \). The array response can be written as

\[
y = A(\theta)s + n,
\]

where \( A(\theta) = [a(\theta_1), \ldots, a(\theta_K)] \) is the virtual array steering matrix with

\[
a(\theta_k) = [1, e^{j(2\pi/\lambda)\sin(\theta_1)}, \ldots, e^{j(2\pi/\lambda)(M_t M_r - 1)\sin(\theta_1)}]^T.
\]

Here, \( n \) is a noise term, and \( s = [\beta_1, \ldots, \beta_K]^T \), where \( \beta_k \) denotes the target reflection coefficient for the \( k \)th target. The array response at a particular time instance consisting of data obtained at all the virtual receivers and corresponding to the same range–Doppler bin is defined as the array snapshot. In highly dynamic automotive scenarios, usually only a small number of array snapshots, or even a single snapshot in the worst case, is available [47].

---

**IEEE Signal Processing Magazine** | July 2020 | 105
In automotive MIMO radar with a virtual ULA, angle finding can be done with digital beamforming (DBF) [12, 48, 49] by performing FFTs on snapshots taken across the array elements, i.e., y in (3) (see Figure 9). DBF can be implemented efficiently in an embedded DSP with a single snapshot. However, DBF is not a high-resolution angle-finding method. Higher-resolution angle finding can be achieved with subspace-based methods, such as multiple signal classification (MUSIC) [50] and estimation of signal parameters via rational invariance techniques (ESPRIT) [51–54], sparse sensing-based methods [55–63], and the iterative adaptive approach (IAA) of [64] and [65]. The performance of subspace-based angle-finding methods relies on accurate estimation of the array covariance matrix with multiple snapshots, which is a challenging task in the highly nonstationary automotive radar scenarios. In such a context, spatial smoothing [66] is applied for introducing virtual snapshots for array covariance-matrix estimation. While sparse sensing-based methods and IAA have a high computational cost, they yield angle estimates based on a single snapshot, which is important for snapshot-limited automotive radar.

Achieving a high angular resolution for the L4 and L5 autonomous driving requirement using a ULA with \( d = \frac{\lambda}{2} \) is very expensive. According to [24], the 3-dB beamwidth of an antenna array with aperture size \( D \) is \( \delta_a = 2 \arcsin(1.4\lambda/\pi D) \). To achieve a 3-dB beamwidth of 1°, the antenna-array aperture should be roughly \( D = 51\lambda \). If the antenna array is a ULA with its interelement spacing as one-half of a wavelength, it should be composed of approximately 100 array elements. Even with the help of MIMO radar technology, the cost of synthesizing such a large virtual ULA with half-wavelength element spacing is very high. One way to further reduce the cost without sacrificing the high angular resolution is via the use of nonuniform, or sparse linear arrays (SLAs) [67–72] synthesized with MIMO radar technology. In that context, selecting the locations of the array elements and carrying out angle finding with the virtual sparse array are key problems.

**High-resolution angle finding with ULAs**

Subspace methods with spatial smoothing

The performance of subspace-based angle finding methods requires an estimate of the array covariance matrix. Such an estimate is typically obtained based on multiple snapshots. However, in the highly dynamic automotive environment, it is not possible to obtain enough snapshots before the model of (3) changes. In such scenarios, spatial smoothing [66] can introduce virtual snapshots for array covariance-matrix estimation. In spatial smoothing, the array snapshot, \( y \), is divided into overlapped subarrays of length \( L \), and a new sampled array covariance matrix \( R \in \mathbb{C}^{L \times L} \) is obtained based on the subarray snapshots.

The eigenvalue decomposition of \( R \), along with the Akaike information criteria metric [73] or the minimum description length metric [74], can be used to identify the number of targets. It should be noted, however, that many ideal assumptions in the deduction of these criteria (including additive white Gaussian noise that is uncorrelated with the source signal) and the availability of enough snapshots for an accurate covariance matrix estimation might not be satisfied in practice. The target angles can be found by identifying the locations of peaks of the MUSIC pseudospectrum [50], \( P(\theta) \), computed at all possible \( \theta \)s, i.e.,

\[
P(\theta) = \frac{1}{a^H(\theta) U^n U^H_n a_1(\theta)},
\]

where \( U_n \) is the noise subspace of \( R \) and \( a_1(\theta) \) is the array steering vector of length \( L \) corresponding to search direction \( \theta \). The computation cost of the MUSIC algorithm is high due to the angle search process. Alternatively, the ESPRIT algorithm could be used for angle estimation [51].
ESPRI T is also a subspace method, which exploits the array shift-invariance properties, and has been widely used in practice. It has a lower complexity than MUSIC, which, however, comes at the cost of reduced angular resolution. ESPRIT requires $2L$ sensors, with $L > K$. To achieve the same angular resolution as MUSIC, ESPRIT needs twice as many sensors. Since 2D arrays are needed in automotive radar to estimate both the azimuth and elevation angles, 2D ESPRIT algorithms [54] can be applied if the array element spacing is uniform rectangular.

Compressive sensing

MIMO radars equipped with mm-wave technology offer a wide bandwidth, thus achieving a high range resolution. As a result, there are only a small number of targets that fall in the same range–Doppler bin, and thus the targets are sparse in the DoA space. This property can be exploited by sparse sensing-based high-resolution methods for target angle estimation. To apply compressive sensing for DoA estimation, the whole DoA FoV is discretized into a fine grid. Assume that the DoA space is discretized on a fine grid with $N$ points and that there are $K$ targets on the grid. The array response in (3) can be rewritten as

$$\mathbf{y} = \mathbf{A} \mathbf{x} + \mathbf{n},$$

where $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ is the basis matrix, with $\mathbf{a}(\theta)$ denoting the array steering vector corresponding to the $i$th grid point, and $\mathbf{x} = [\beta_1, \beta_2, \ldots, \beta_K]^T$ is a sparse vector with $K$ nonzero elements. The value of $\beta_i$ is nonzero if there is a target at the $i$th grid point. The coherence of the basis matrix, defined as

$$\mu \triangleq \max_{i \neq j} \frac{|\mathbf{a}(\theta_i)^H \mathbf{a}(\theta_j)|}{\|\mathbf{a}(\theta_i)\|_2 \|\mathbf{a}(\theta_j)\|_2},$$

needs to be low for obtaining uniform recovery guarantees [75]. When meeting the required coherence conditions, the DoA can be found by solving an $\ell_1$-norm optimization problem, such as the Dantzig selector [76], defined as

$$\min \|\mathbf{x}\|_{\ell_1} \quad \text{s.t.} \quad \|\mathbf{A}^H (\mathbf{y} - \mathbf{A} \mathbf{x})\|_2 < \eta.$$  

or greedy methods, such as orthogonal matching pursuit (OMP) [77].

In the preceding formulation, targets are assumed to be on the grid, which is not always possible in practice. While one can make the grid finer to capture the targets, the coherence of matrix $\mathbf{A}$ would increase, which would make the $\ell_1$-norm solution invalid [78]. Thus, the performance of compressive sensing-based methods is sensitive to targets appearing off the grid [79]. Sparse sensing and matrix completion-based methods [62], [63] can avoid grid issues without sacrificing the high-resolution performance.

IAA

The covariance matrix of $M$ array snapshots $\mathbf{y}_l$, $l = 1, \ldots, M$, can be written as $\mathbf{R} = \mathbf{A}(\theta) \mathbf{P} \mathbf{A}(\theta)^H$, where $\mathbf{P}$ is a $K \times K$ diagonal matrix whose diagonal elements contain the power of target reflections. Angle finding in the IAA algorithm [64], [65] is carried out by iteratively estimating the reflection coefficient $\beta_i$. The estimate is found by minimizing the weighted least-square cost function $\sum_{l=1}^{M} \| \mathbf{y}_l - \beta_i(l) \mathbf{a}(\theta_i) \|^2_{\ell_0} / (\gamma_0)$, where $\| \mathbf{x} \|^2_{\ell_0} = \sum_i |x_i|^2$. The interference and noise covariance matrix $\mathbf{Q}(\theta_i) = \mathbf{R} - \hat{P}_i \mathbf{a}(\theta_i) \mathbf{a}(\theta_i)^H$. The solution is given by [64]:

$$\hat{\beta}_i(l) = \frac{\mathbf{a}(\theta_i)^H \mathbf{R}^{-1} \mathbf{y}_l - \mathbf{a}(\theta_i)^H \hat{\mathbf{P}}_i \mathbf{a}(\theta_i)}{\mathbf{a}(\theta_i)^H \mathbf{Q}^{-1} \mathbf{a}(\theta_i)},$$

(9)

Then matrix $\mathbf{P}$ can be updated as $\hat{\mathbf{P}}_i = (1/M) \sum_{l=1}^{M} |\hat{\beta}_i(l)|^2$. In IAA algorithm implementation, the DoA space is discretized into a fine grid of $N$ points, and steering matrix $\mathbf{A}$ is constructed in the same way as in compressive sensing. In addition, a standard delay-and-sum beamformer is used to initialize $\mathbf{P}$:

$$\hat{\mathbf{P}}_i = \frac{\sum_{l=1}^{M} |\mathbf{a}(\theta_i)^H \mathbf{y}_l|^2}{M |\mathbf{a}(\theta_i)^H \mathbf{a}(\theta_i)|^2}.$$  

(10)

High-resolution angle finding with SLAs

As stated before, the cost of synthesizing a large virtual ULA of $D$ elements with half-wavelength element spacing is very high. One way to further reduce the cost without sacrificing the high angular resolution is via the use of nonuniform or SLAs [67], [68], [80]. With MIMO radar technology, $M, M < D$ virtual array elements can be synthesized. To make the SLA aperture the same as the ULA, two virtual-array elements should be deployed at the edge locations of the ULA. For the remaining virtual array elements, there are multiple possibilities to deploy. The main issue with the SLA is that the grating lobes may introduce ambiguity into angle finding. In that context, the key problems are how to select the locations of the array elements such that the peak sidelobe level (PSL) of the virtual SLA beampattern is low and how to carry out angle finding. There is no analytical solution to determining the antenna locations that achieve a minimum PSL for a given number of antennas [81]. Optimal sparse array design requires global optimization techniques, such as particle-swarm optimization [68], [82], [83].

In automotive MIMO radar with a virtual SLA, angle finding can still be done with conventional FFT or ESPRIT methods if the holes in the virtual SLA can be filled via interpolation or extrapolation techniques to mitigate the grating lobes [16], [84]. Alternatively, instead of filling the holes, angle finding of the sparse array can be done using spatial compressive sensing ideas [85]. In the SLA scenario, it can be easily verified that the coherence of the basis matrix [see (7)] is the PSL of the SLA array beampattern [67]. Therefore, the coherence (or, equivalently, the PSL) of a sparse array plays a key role in obtaining uniform recovery guarantees for compressive sensing [75]. If the PSL of the SLAs is low, angle finding using SLAs can be done via compressive sensing or IAA.
In Figure 10, we give an example of a virtual SLA with an aperture of 19λ, synthesized with MIMO radar technology using four transmit and four receive antennas. The first and fourth transmit/receive antennas are deployed at the edge of the physical aperture, while the remaining antennas are chosen such that the PSL is $-9.1$ dB. Angle estimation via the IAA when using the sparse linear array of Figure 10 is illustrated in Figure 11. The ground truth involves two targets with azimuths of 5 and 10°. The SNR of the received beam vector is set to 30 dB. For comparison, the FFT spectrum is also plotted. One can observe the sharper peaks around the target azimuth angles and the more attenuated sidelobe in the IAA spectrum as compared with the FFT spectrum.

**High-resolution imaging radar for autonomous driving**

Automotive radar with a small number of antennas has been used for ADAS purposes since the late 1990s. Such radar mostly sees point targets and obtains velocity information. However, the current generation of automotive radar for ADAS has a rather limited ability to resolve closely spaced targets. Lidar systems have better angular resolution (less than 1°) and have been introduced into L4/L5 autonomous driving systems. Lidar can provide point clouds. Through the use of deep neural networks, such as PointNet [86] and PointNet++ [87], the point clouds can lead to target identification. However, due to its use of the light-spectrum wavelength, lidar is susceptible to bad weather conditions, such as fog, rain, snow, and dust in the air. In addition, the cost of lidar is high. On the other hand, automotive radar with mm-waveform technology has the potential to provide point clouds at a much lower cost than lidar and with more robustness to weather conditions. Such radar is referred to as *high-end radar* or *imaging radar* [88]. Computer vision techniques [86], [87] that were previously reserved for high-resolution camera sensors and lidar systems can be applied to imaging-radar data to identify targets. For example, a car can be identified based on the 2D radar points of an imaging radar using PointNet [89]. Imaging radar has been attracting the interest of major Tier-1 suppliers and automotive radar start-ups.

MIMO radar is a good candidate for high-resolution imaging radar for autonomous driving. In MIMO radar using FMCW waveforms, the targets are first distinguished in the range and Doppler domains. Then, a large virtual array with hundreds of elements can be synthesized to provide a high resolution in both the azimuth and elevation. As a result, point clouds with performance similar to lidar can be generated at a much lower cost. In this section, we introduce the concept of imaging radars using MIMO technology; present some examples for synthesizing hundreds of virtual array elements by cascading multiple radar transceivers, with each supporting a small number of antennas; and discuss design challenges.

**Cascade of multiple radar transceivers**

Today, most of the automotive radar transceivers designed for ADAS functionality, such as the MR3003 from NXP Semiconductor and AWR1243 from Texas Instruments, can support up to three transmit and four receive antennas. Therefore, using a single automotive radar transceiver with MIMO radar technology, only 12 virtual array elements can be synthesized. To meet the requirement for L4 and L5 autonomous driving, multiple automotive radar transceivers would need to be cascaded together, with all the transceivers synchronized as a single...
unit. The received data from all the receive antennas would be processed coherently. Cascading provides a cost-effective and scalable solution to achieve a high angular resolution.

In [90], General Motors and Texas Instruments successfully demonstrated that up to four Texas Instruments AWR1243 radar chips can be cascaded to provide 12 transmit and 16 receive antennas, enabling a synthesis of 192 virtual array elements. In [91], a prototype of five cascading Infineon radar chips was built to synthesize a virtual array of 128 × 4 elements. Such a high number of virtual array elements provides a lot of opportunities in array design. Several azimuth and elevation array configurations can be found in [16]. Usually, a trade-off between balancing the angular resolution in the azimuth and elevation needs to be considered.

Examples of cascaded imaging radars

Figure 12 shows an imaging radar design reference board that has 12 transmit and 16 receive antennas, formed by cascading four Texas Instruments AWR1243 radar transceivers [92]. The azimuth FoV is [−70°, 70°]. One transceiver is selected as the master and all the others as slaves for the clock distribution. In this way, synchronization can be achieved among four transceivers, enabling coherent FMCW transmission from the 12 transmit antennas and joint data processing from the 16 receive antennas. The array configuration of cascaded imaging radar is illustrated in Figure 12. There are three transmit antennas placed along the vertical direction for elevation angle finding and nine transmit antennas placed along the horizontal direction for azimuth angle finding. The virtual array in the horizontal direction is a dense ULA with half-wavelength spacing, and it consists of 86 virtual array elements (the overlapped virtual array elements are not shown). The array aperture in the azimuth direction is \( D_A = 42.5 \lambda \). In antenna theory, the 3-dB beamwidth defines the angular resolution. According to [24], the 3-dB beamwidth of the azimuth angle is

\[
\Delta \theta_A = 2 \arcsin \left( \frac{1.44 \lambda}{\pi D_A} \right) \approx 1.2^\circ. \tag{11}
\]

In the vertical direction, the antennas in the three elevation positions form multiple minimum redundancy arrays (MRAs) [93] along the horizontal direction. Angle finding in the MRAs requires multiple snapshots. These MRAs along the horizontal direction can be used as snapshots for elevation angle finding. The elevation array aperture is \( D_E = 3 \lambda \), and the 3-dB beamwidth of elevation is

\[
\Delta \theta_E = 2 \arcsin \left( \frac{1.44 \lambda}{\pi D_E} \right) \approx 17^\circ. \tag{12}
\]

In the second example [94], an imaging radar tested with \( M_T = 24 \), \( M_R = 24 \) antennas using a TDM scheme (see Figure 13) is presented. The virtual array apertures have been doubled in both the azimuth and elevation directions. The missing elements in the vertical direction in the middle have been interpolated [84]. After filling the holes and removing the redundant elements, the virtual array is a uniform rectangular array with 25 × 23 = 575 elements. Wider FoVs of [−25°, 25°] in both the azimuth and elevation were considered in [94]. However, for the forward-looking LRR sensors, the typical azimuth and elevation FoVs are [−15°, 15°] and [−5°, 5°], respectively [5]. Consequently, the array interelement spacing in horizontal and vertical directions can be set to \( d_x = 1.93 \lambda \) and \( d_y = 5.73 \lambda \), respectively. Since the interelement spacing is larger than one half wavelength, there are grating lobes in both the azimuth and elevation. However, the grating lobes out of the FoVs can be suppressed through the antenna element design. If the carrier frequency is \( f_c = 77 \text{ GHz} \), the real size of the physical 2D antenna array is roughly 10 × 25 cm. Then, the 3-dB beamwidth of the azimuth and elevation beampattern is [24]:

\[
\Delta \theta_A = 2 \arcsin \left( \frac{1.44 \lambda}{\pi 24 d_x} \right) \approx 1.1^\circ, \tag{13}
\]

\[
\Delta \theta_E = 2 \arcsin \left( \frac{1.44 \lambda}{\pi 22 d_y} \right) \approx 0.4^\circ. \tag{14}
\]

It is worth noting that the mutual coupling between array elements is reduced significantly if the interelement spacing is larger than one half wavelength, which reduces the burden
for the array calibration [94]. If we want to keep the azimuth FoV unchanged and increase the azimuth angular resolution to roughly 0.4°, each horizontal receive array needs to add 21 more antennas; i.e., the dimension of the new physical antenna array is 66 × 24 with size of approximately 25 × 25 cm. The angular resolution of 0.4° in both the azimuth and elevation directions provided by imaging radar is similar to the azimuth and elevation resolution of Velodyne’s lidar products, such as the HDL-64E, with a spin rate of 20 Hz [95].

**Design challenges of imaging radar**

Achieving waveform orthogonality in imaging radars using FMCW with a large number of transmit antennas is quite challenging. One strategy could be to divide the transmit antennas into several subgroups. In each subgroup, the transmit antennas would transmit simultaneously with slow time phase coding (DDM), while antennas of different subgroups would be scheduled to transmit in different time slots (TDM).

Clock distribution among multiple cascaded transceivers is also challenging. For FMCW mixer operation, an LO is shared among the master and slaves, and the LO routing from the master to all the slaves in the circuit should be matched. Also, the additional ADC sampling and data transmission among different transceivers needs to be synchronized. It is desirable to develop an automotive radar transceiver that can incorporate a large number of transmit and receive antennas. For example, Uhnder has developed a radar system-on-chip (SoC) that has 12 transmit and 16 receive antennas, enabling the synthesis of 192 virtual array elements [96]. Thus, the four current automotive radar transceivers in the cascaded imaging radar shown in Figure 12 can be replaced with a single SoC radar chip. The radar on chip developed by Vayyar has 48 transceivers at 76–81 GHz, which can provide synthesis across 2,000 virtual-array elements [97].

**Challenges in automotive MIMO radar**

In this section, we discuss signal processing challenges that cannot be addressed by the straightforward application of existing ideas and that could inspire new research.

**Angle finding in the presence of multipath reflections**

Automotive radar runs in multipath scenarios [98]. In general, radio propagation in the presence of multipath occurs along four possible routes, i.e., direct/direct, direct/indirect, indirect/direct, and indirect/indirect routes. Figure 14 shows a vertical multipath scenario where the height of the radar and target are h3 and h1, respectively. The length of the direct/direct path is d1 = 2r1; the length of the direct/indirect or indirect/direct path is d2 = d3 = r1 + r2 + r3; and the length of the indirect/indirect path is d4 = 2(r2 + r3). The received signal, having gone through the four paths, can be written as

\[ y_r = \sum_{i=1}^{4} \beta_i e^{j2\pi \frac{r_i}{\lambda}}, \quad (15) \]

where the amplitude term \( \beta_i \) is a function of the antenna gain, path loss, road-reflection coefficient, and target RCS. The signal from the indirect paths and that from the direct path may arrive out of phase and thus add up destructively. As a result, the power of the received signal would fluctuate with distance [98], and thus angle finding at SNR nulls would be unstable.

MIMO radar with colocated transmit and receive antennas, also referred to as monostatic MIMO radar, is based on the assumption that the DoD and DoA are equal. However, in the presence of multipath, that assumption does not hold, and the system becomes bistatic [99], [100]; i.e., the transmit and receive antennas view the target from different aspect angles. Figure 15 shows a vehicle moving parallel to the guardrail, with an SRR sensor mounted at its front left corner. The length of the direct path of the radar signal is d1 = 2r1, corresponding to \( \theta_i = \theta_r = \theta_t \). There are also multipath reflections due to the guardrail. The range of the first multipath reflection is d2 = 2(r1 + r2 + r3), corresponding to \( \theta_i = \theta_r, \theta_r = \theta_t \) or \( \theta_i = \theta_2, \theta_r = \theta_1 \). The range of the second multipath reflection is d3 = 2(r2 + r3), corresponding to \( \theta_i = \theta_t = \theta_2 \). Compared to the direct path, multipath reflections result in longer-range and smaller Doppler. For the first type of multipath, the range and Doppler bin is the same as mirror-image target detection. However, as \( \theta_i \neq \theta_t \), it turns out that the phase of each virtual array element is corrupted. In other words, the monostatic MIMO radar assumption does not hold, which results in a “ghost” target whose direction is different from the mirror target.

To solve this issue, some ideas have been proposed in [99]–[101]. For example, joint estimation of the DoD and DoA is proposed.

---

**FIGURE 13.** An example of imaging radar using MIMO radar technology with Mh = 24, Mr = 24. (a) The physical antenna configuration. (b) The virtual antenna array.

**FIGURE 14.** The direct path and the vertical multipath [98].
in [99]; however, by ignoring the structure of the transmit array, that method is not able to enjoy the benefit of the synthesized virtual array. Polarimetric features are exploited in [100] to separate objects in a multipath scenario. However, the approach in [100] can separate only certain real target cases from their mirror targets; for example, when the real target is known and the polarization state change of the multipaths can be recognized. The method in [100] does not work when the ghost target direction is different than that of the mirror target. Doppler information can be exploited to detect moving vehicles in urban areas under multipath [101]. However, Doppler information is not always available when both the objects and host vehicles are stationary. In general, there is a need for more research addressing the ghost target issue in MIMO radar due to multipath.

Waveform orthogonality in automotive MIMO radar

As stated in the “Introduction of Automotive Radar With MIMO Radar Technology” section, different strategies, such as TDM, DDM, and FDM, can be adopted in automotive FMCW radars to achieve waveform orthogonality. However, several challenges associated with each strategy need to be addressed. For example, in the TDM scheme, the scheduling delay between transmit antennas may introduce phase error for a moving target, which needs to be compensated for; otherwise, the synthesized array beampattern will be distorted. Further, the maximum unambiguous detectable velocity under TDM is reduced by a factor of $M_t$. In the DDM scheme, the Doppler sidelobes are high due to the residual of the phase coding. As a result, targets with small RCSs, e.g., pedestrians, that are close to the target with strong reflections, e.g., trucks, might be masked. The search time for phase codes using stochastic algorithms increases exponentially as the code length increases. Computationally efficient algorithms are needed to address this problem. In the FDM scheme, although a randomization of the frequency shift among transmit antennas could reduce the range-angle coupling, a large number of transmit antennas would be needed for the improvement to be notable [41].

Recently, PMCW has been proposed for achieving orthogonality [102]–[105]. Each antenna transmits a sequence of phase-coded pulses. Let $x_m = [x_m(1), \ldots, x_m(N_p)]^T$ be the complex unimodular code sequence of the $m$th transmit antenna, where $x_m(n) = e^{j\phi_m(n)}$ is the $n$th code of $x_m$ and $N_p$ is the code length. Here, the phase $\phi_m(n)$ can be chosen arbitrarily in $[-\pi, \pi]$. The duration of a single code sequence is $T_p = N_p T_c$, with $T_c$ being the duration of a subpulse. In practice, binary code sequences have been widely used due to their simplicity. The bandwidth of PMCW is $B = 1/T_c$. The time–bandwidth product of a code sequence is $BT_p = N_p$. Since the pulses are transmitted continuously, the code sequences should have good periodic autocorrelation and cross-correlation properties [102]. The periodic cross correlation of two code sequences $x_m$ and $x_l$ at lag $k$ is defined as

$$r_{ml}(k) = \sum_{n=1}^{N_p} x_m(n) x_l^* ((n + k) \mod(N_p)).$$  \hfill (16)\

when $m = l$, $r_{ml}(k)$ becomes the periodic autocorrelation function of $x_m$. Good correlation properties require that the values of the periodic autocorrelation at nonzero lag and that the values of the cross correlation at any lag be low. The Welch lower bound on the cross correlation between any pair of binary sequences with a period of $N_p$ in a set of $M$ sequences equals [106]

$$r_{ml}(k) \geq N_p \sqrt{\frac{M_t - 1}{M_t N_p - 1}} \approx N_p. \quad (17)$$

Good periodic cross-correlation properties help achieve waveform orthogonality, while good periodic autocorrelation properties make it easier to use matched filters to extract signals reflected from the range bin of interest and suppress signals reflected from other range bins.

As compared to FMCW, PMCW radar has several advantages. PMCW radar is better suited for achieving waveform orthogonality in imaging radars with a large of number of transmit antennas. PMCW radar can take advantage of existing sequences with good autocorrelation and cross-correlation properties that were previously developed for code-division multiple accessing (CDMA) communications, such as Gold, Kasami, and m-sequences [107]–[109]. Further, in PMCW radar, each automotive radar sensor can have a unique digital sequence, which may help reduce the automotive radar mutual interference. As a bonus, PMCW radar also provides certain communication capability [110] and thus can be explored as a dual-functional radar communication system [111].

However, PMCW radar has many implementation challenges. First, the sampling rate of the ADC should satisfy

**FIGURE 15.** A typical multipath reflection scenario along a guardrail for an SRR sensor mounted at the vehicle’s front left corner.
the Nyquist rule; i.e., \( f_s \geq 2B = 2/T_\cdot \). The high bandwidth required for a high range resolution necessitates a high-speed ADC and high-speed processing hardware. In practice, keeping the resolution of the ADC as low as possible is required \[112\]. Second, according to the Welch bound of (17), the cross-correlation lower bound of any pair of binary sequences is of the order of \( \mathcal{O}(\sqrt{N_\rho}) \), which might not provide a sufficient separation of the transmit waveforms of different antennas. In practice, the autocorrelation and cross correlation of the code sequences are desired to have low sidelobes within a low correlation zone. Furthermore, because there is no mapping relationship between the range and beat signal in PMCW, it would be difficult to use high-pass analog filters to reject or attenuate ultraclose-range return signals, including direct path signals from the transmit antennas, reflections from the radome, and vehicle bumpers. This escalates the dynamic-range challenge, especially when the resolution of the ADC must be kept as low as possible \[112\].

**Mutual-interference mitigation**

Automotive radar mutual interference is a challenging issue that needs to be addressed. The use of radar for ADAS and autonomous driving is climbing rapidly. As the number of vehicles equipped with automotive radar is escalating, with each vehicle deploying up to 10 automotive radar units, the probability of mutual interference between automotive radar units increases. Figure 16 shows an example of an automotive radar interference scenario where two front-looking automotive radar sensors from two stopped cars illuminate each other. If the radar sensors operate at the same frequency band and transmit at the same time, they will interfere with each other. Without interference mitigation, automotive radar sensors suffer from performance degradation \[113\]. Therefore, it is desired to optimize both the radar transmit- and receive side operations to mitigate the interference.

A typical mutual interference scenario for two automotive FMCW radars is given in Figure 17. The blue line indicates the chirp of an FMCW radar mounted on the host vehicle, with a faster sweep rate \( S_1 \), referred to as the *victim radar*, while the red line denotes the chirp of an FMCW radar mounted on another vehicle, with a slower sweep rate \( S_2 \), referred to as the *interference radar*. The interference radar is set to illuminate the FoV of the victim radar directly. At the receiver of the victim radar, the duration of the corrupted samples introduced by interference in one pulse is \( T_{int} = 2\left( f_s^{max} / (S_1 - S_2) \right) \), where \( f_s^{max} \) is the highest cutoff frequency of the antialiasing BPF \[114\]. After mixing with the transmitted chirp, in addition to the beat frequency corresponding to real targets, the corrupted samples contain frequencies spanning the whole interval of

![Figure 16](image1.png)

**Figure 16.** An illustrative automotive radar interference scenario, where two front-looking automotive radar sensors illuminate each other.

![Figure 17](image2.png)

**Figure 17.** (a) Automotive radar interference exists when the victim and interference radars have different FMCW sweep rates. (b) After down-conversion and passing through the BPF, the sequence sampled from one chirp is produced. (c) Since corrupted samples introduced by interference have a much larger amplitude than the samples of the target echo, they can be clipped or gated.
the BPF (see Figure 17). Further, since the interference signal arrives via a direct path, it is much stronger than the echoes of the target.

Intuitively, this type of corrupted samples introduces extra energy that spans the whole beat frequency band of interest. As a result, after performing range and Doppler FFTs on the receive signal with interference, the noise floor in the range–Doppler spectrum would increase, as shown in Figure 18. As a result, targets with a small RCS might be masked. For example, the pedestrian in Figure 16 might be buried in the noise and may not be detected by automotive radar when the interference is strong.

The amount of interference energy can be viewed as an integral of corrupted samples through the interference duration interval $T_{int}$. Therefore, the interference can be greatly suppressed by clipping, or gating, the corrupted samples. In particular, the corrupted samples with an amplitude larger than a threshold close to the amplitude of the target echo plus noise would be clipped (see Figure 17). However, gating does not completely remove the interference because the clipped corrupted samples still contain frequency components that occupy the whole beat frequency band of interest.

Figure 18 presents the simulated range and Doppler spectra of victim radar. Three scenarios have been simulated: target plus noise only, target plus noise plus interference, and target plus noise plus interference with gating. It can be seen that without gating, the target is masked by the increased noise that is due to the interference. Fast time-domain gating helps to recover the target in the range and Doppler domains. It should be noted that this type of interference needs to be mitigated in fast time after ADC sampling, and thus the mitigation algorithms not only need to be effective but also efficient. Gating has a low computational cost, and it can significantly suppress the noise level in the range and Doppler spectra.

Other interference-mitigation approaches include repairing the corrupted samples in the time domain, frequency domain [115], or spatial domain using adaptive beamforming techniques [116]. However, if the FMCW sweep rates of the interference and victim radar are the same, the interference radar will create ghost targets [117], which makes interference mitigation more challenging, since this scenario is difficult to detect.

With the introduction of PMCW radar, the mutual interference between automotive radar sensors can be greatly mitigated in the code domain. The PMCW code sequences used in different automotive radar sensors are typically not time aligned. Therefore, to suppress the mutual interference, the periodic cross correlation of any pair of sequences should be low at all lags. PMCW radar is similar to CDMA in communication systems. In other words, the interference will be a

| Table 1. The different DoA estimation algorithms in automotive radar scenarios. |
|-----------------|---------|---------|--------|--------|--------|--------|--------|
| Algorithm       | Resolution | Snapshot | Array   | Grid-Free | Rank Estimation | Robustness | Complexity |
| DBF             | Low      | Single   | ULA/SLA | No       | No      | Strong  | Low     |
| MUSIC           | High     | Multiple | ULA     | No       | Yes     | Medium  | High    |
| ESPRIT          | High     | Multiple | ULA     | Yes      | Yes     | Medium  | Medium  |
| OMP             | High     | Single   | ULA/SLA | No       | No      | Medium  | High    |
| IAA             | High     | Single   | ULA/SLA | No       | No      | Strong  | High    |

Machine learning algorithms and deep neural networks have also been applied in automotive radar for target recognition and classification.
wide-band pseudonoise signal. However, the mutual interference mitigation using PMCW highly depends on the periodic cross-correlation properties of the code sequence. Therefore, designing code sequences with good periodic autocorrelation and cross-correlation properties is of great interest. The other research problems include investigating the interference between FMCW and PMCW radars [118].

**Efficient, high-resolution angle finding algorithms are needed**

A typical duration of a look in automotive radar is roughly 50 ms, corresponding to a detection update rate of 20 Hz [119]. In such a short duration, the current generation of automotive radar for ADAS can report a maximum of 64–200 detections. With high-resolution imaging radar, the number of cells that can be selected for angle finding from the 2D range and Doppler spectrum is approximately 10,000 in a single look for a typical road scenario [15]. To achieve point clouds for autonomous driving, angle finding needs to be performed thousands of times in a single look, which is a great challenge for imaging radar with hundreds of virtual array elements. Computationally efficient, high-resolution angle finding algorithms are highly desirable for real-time implementation in automotive radar.

To reduce the computation complexity, beamspace ESPRIT [52] and unitary ESPRIT [53] algorithms have been proposed. The idea of beamspace ESPRIT is to decompose the original ULA vector into several low-dimensional beamspaces via a transform, such as the FFT. Then, if the beamspace transform matrix has the same shift-invariance structure, angle finding can be carried out via ESPRIT on each beamspace in parallel, with a reduced computational time [52]. The unitary ESPRIT algorithm takes advantage of the unit-magnitude property of the phase factors representing the phase delays between the two subarrays and is formulated in terms of real-valued computations. As a result, it achieves a great reduction of the computational complexity [53].

The computation cost of each IAA iteration is \(2NM^2 + NM + M^3\), where \(M\) is the number of array snapshots and \(N\) is the number of discretized grids. Fast and superfast IAA algorithms have been proposed in [120]–[122], respectively. The fast IAA algorithm exploits the FFT operation as well as the Gohberg–Semencul representation of matrix \(\mathbf{R}^{-1}\). As a result, the computation cost of each fast IAA iteration is \(M^2 + 12\zeta(2M) + 3\zeta(N)\), where \(\zeta(N)\) stands for the computation cost of performing an FFT of size \(N\), i.e., \(O(N \log N)\) [121]. The superfast IAA uses a conjugate-gradient algorithm to approximate the matrix \(\mathbf{R}^{-1}\), which further reduces the computation cost.

The strengths and limitations of each DoA estimation algorithm discussed in the “Angle Finding in Automotive MIMO Radar” section when they are applied to the automotive radar scenario are summarized in Table 1. For subspace-based, high-resolution DoA estimation methods, such as MUSIC and ESPRIT, the automotive radar array needs to be a ULA, and multiple snapshots are required to estimate the array covariance matrix accurately. However, automotive radar operating in a highly dynamic environment typically relies on a single snapshot. While multiple snapshots can be generated via spatial smoothing or by dividing a chirp into subchirps, the associated cost is a respectively reduced array aperture or reduced SNR. SLAs have been widely used in automotive radar to further reduce the hardware cost. However, it is not straightforward to apply MUSIC or ESPRIT to SLAs-based automotive radar.

On the other hand, DBF and sparsity-based, high-resolution methods, such as the OMP and IAA, apply to SLA-based as well as ULA-based automotive radar and work with a single snapshot. In the DBF method, the number of targets can be estimated by counting the number of peaks in the DoA spectrum. DBF is not sensitive to coherent or correlated signals, which in subspace-based methods [123] need special preprocessing via spatial smoothing. It has been shown that DBF is robust to array-element position errors and has a low computational cost [123]. However, DBF is not a high-resolution method. Also, since the OMP and IAA are iterative schemes, they involve a high computation cost, which limits their applicability in the low-cost embedded DSPs typically used in current-generation automotive radar. Further, the methods of DBF, MUSIC, OMP, and IAA assume targets are on the grid and suffer from errors when the targets arise between grid points. In summary, more research is needed on developing computationally efficient, high-resolution DoA estimation algorithms that are robust to noise and applicable to automotive radar using SLAs with a low PSL under a single snapshot.

**Summary**

We have reviewed the advantages of MIMO radar technology in increasing the angular resolution of commercial automotive radar while, at the same time, offering a low hardware cost and small package size. In particular, we have shown that MIMO radar technology plays a key role in high-resolution imaging-radar systems for L4 and L5 autonomous driving. Automotive MIMO radar technologies, such as waveform orthogonality with slow time phase coding, time-division multiple access, and non-uniform linear array synthesis with low-peak sidelobes have been discussed. High-resolution angle finding literature for snapshot-limited automotive MIMO radar has been reviewed. In particular, high-resolution angle finding methods, including subspace methods with spatial smoothing, compressive sensing, and the IAA, have been discussed.

In particular, we have reviewed methods of cascading multiple automotive radar transceivers to synthesize hundreds or thousands of virtual array elements in high-resolution imaging radar systems and have discussed related design challenges. We have also addressed the challenges of MIMO radar technology in automotive applications, which would inspire further research for the signal processing community. In addition, we
have discussed angle finding issues in the presence of multipath reflections, waveform orthogonality strategies, automotive mutual interference mitigation, and computationally efficient, high-resolution angle finding methods.

Authors

Shunqiao Sun (shunqiao.sun@ua.edu) received his Ph.D. degree in electrical and computer engineering from Rutgers University, New Jersey in 2016. In 2019, he joined the Department of Electrical and Computer Engineering at the University of Alabama, Tuscaloosa, as a tenure-track assistant professor. His research interests lie at the interface between statistical and sparse signal processing with mathematical optimizations, multiple-input, multiple-output (MIMO) radar, automotive radar, remote sensing, machine learning, and connected and autonomous vehicles. He was awarded the 2015–2016 Rutgers University Electrical and Computer Engineering Graduate Program Academic Achievement Award. He was also the winner of the 2016 IEEE Aerospace and Electronic Systems Society Robert T. Hill Best Dissertation Award for his thesis “MIMO Radar With Sparse Sensing.” He is a Senior Member of the IEEE.

Athina P. Petropulu (athinap@rutgers.edu) received her B.S. degree from the National Technical University of Athens, Greece, and her M.S. and Ph.D. degrees from Northeastern University, Boston, all in electrical and computer engineering. She is a distinguished professor in the Department of Electrical and Computer Engineering, Rutgers University, New Jersey, which she chaired from 2010 to 2016. Her research interests include statistical signal processing, wireless communications, signal processing in networking, physical-layer security, and radar signal processing. She received the 1995 Presidential Faculty Fellow Award from the National Science Foundation and the White House. In 2005, she received the IEEE Signal Processing Magazine Best Paper Award, and in 2012, the IEEE Signal Processing Society Meritorious Service Award for exemplary service in technical leadership capacities. She is the president-elect of the IEEE Signal Processing Society and a Fellow of the IEEE.

H. Vincent Poor (poor@princeton.edu) received his Ph.D. degree in electrical engineering and computer science from Princeton University, New Jersey in 1977. He is the Michael Henry Strater University Professor of Electrical Engineering at Princeton, where he served as dean of the School of Engineering and Applied Science from 2006 to 2016. His research interests include information theory and signal processing and their applications in wireless networks, energy systems, and related fields. He received the IEEE Signal Processing Society Technical Achievement and Society Awards in 2007 and 2011, respectively; the 2017 IEEE Alexander Graham Bell Medal; the 2019 American Society for Engineering Education Benjamin Garver Lamme Award; a D.Sc. honoris causa from Syracuse University, New York, in 2017; and a D.Eng. honoris causa from the University of Waterloo, Ontario, in 2019. He is a Fellow of the IEEE.

References


Impact of ADC clipping and quantization on phase-modulated 79 GHz CMOS
Recent evolution of automotive imaging
A portable 3-D imaging FMCW MIMO radar demonstrator
Almost perfect autocorrelation sequences for binary phase-modulated continuous wave radar
Minimum-redundancy linear arrays
PointNet: Deep learning on point cloud data with PointNets
PointNet++: Deep hierarchical feature learning using a fully polarimetric automotive radar
MIMO radar system-on-chip
Interference of FMCW and PMCW automotive radars
Impact on phase-modulated 79 GHz MIMO radar transceivers
Wideband 120 GHz to 140 GHz MIMO radar: System design and imaging results
Application of antenna arrays to mobile communications. II.
Continuous wave radar
IAA spectral estimation: Fast implementation using accelerated Prony’s method
Spatial compressive sensing for MIMO radar
The IAA spectral estimate
Coordination of narrowband and wideband CDMA cellular networks
IEEE Radar Conf.
IEEE J. Solid-State Circuits
IEEE J. Solid-State Circuits
IEEE Signal Processing Magazine