FOUR-DIMENSIONAL HIGH-RESOLUTION AUTOMOTIVE RADAR IMAGING
EXPLOITING JOINT SPARSE-FREQUENCY AND SPARSE-ARRAY DESIGN

Shunqiao Sun and Yimin D. Zhang

1Department of Electrical and Computer Engineering, The University of Alabama, Tuscaloosa, AL 35487
2Department of Electrical and Computer Engineering, Temple University, Philadelphia, PA 19122
E-mail: shunqiao.sun@ua.edu, ydzhang@temple.edu

ABSTRACT
We propose a novel automotive radar imaging technique to provide high-resolution information in four dimensions, i.e., range, Doppler, azimuth, and elevation, by exploiting a joint sparsity design in frequency spectrum and array configurations. Random sparse step-frequency waveform is proposed to synthesize a large effective bandwidth and achieve high range resolution profiles. This concept is extended to multi-input multi-output (MIMO) radar by applying phase codes along the slow time to synthesize a two-dimensional (2D) sparse array with a high number of virtual array elements which enable high-resolution direction finding in both azimuth and elevation. The 2D sparse array acts as a sub-Nyquist sampler of the corresponding uniform rectangular array (URA), and the corresponding URA response is recovered by completing a low-rank block Hankel matrix. The proposed imaging radar provides point clouds with a resolution comparable to light detection and ranging (LiDAR) but with a much lower cost and is insensitive to weather conditions.

Index Terms— Automotive radar, autonomous driving, random sparse step-frequency waveform, sparse array, MIMO radar

I. INTRODUCTION
Next-generation automotive radar must provide high resolution estimation in four dimensions, i.e., range, Doppler, and azimuth and elevation angles, yet maintain a low cost for feasible mass production. High-resolution imaging radar is being developed to provide point clouds of the surrounding environment [1]–[4]. Via use of deep neural networks, such as PointNet [5] and PointNet++ [6], point clouds generated by high-resolution imaging radar can lead to adequate target identification. State-of-the-art automotive radars exploit frequency-modulated continuous-waveform (FMCW) signals at millimeter-wave frequencies [4], [7], [8] to enable high-resolution target range and velocity estimation in all weather conditions at a much lower cost than light detection and ranging (LiDAR). To achieve a high range resolution, the transmit signals are designed to occupy a large bandwidth. FMCW radars sweep the frequency linearly over the entire large bandwidth, thereby making the signal susceptible to interference from other automotive radars.

To meet the angular resolution requirement for autonomous driving, the array is required to have a large aperture. However, most of the current automotive radar transceivers designed for advanced driver assistance systems (ADAS) functionality, such as NXP Semiconductor MR3003 and Texas Instruments (TI) AWR1243 [9], only support up to 3 transmit and 4 receive antennas. Such radar units synthesize 12 virtual array elements using the multiple-input multiple-output (MIMO) radar technology. Because of the small array size, such radars are typically used only for azimuth angle estimation. A cost-effective and scalable solution to form a large array is to coherently cascade multiple automotive radar transceivers. For example, up to 4 TI AWR1243 radar chips can be cascaded together to provide 12 transmit and 16 receive antennas, thereby synthesizing 192 virtual array elements [10]. Other cascade products are also reported [11], [12]. Such high number of virtual array elements provide much higher processing capability and enable angle estimation in both azimuth and elevation [3]. Such two-dimensional (2D) angular information in both azimuth and elevation is crucial for autonomous driving. In particular, the height information of targets is important to enable drive-over and drive-under functions.

Conventional radars use uniform linear or rectangular array configurations with half-wavelength interelement spacing. In this case, the aperture is proportional to the number of array sensors. As such, high-resolution target direction-of-arrival (DOA) estimation requires a high number of antennas which are often infeasible for automotive radars due to the strict cost constraints. A cost-effective solution is to use sparse MIMO arrays [13]–[15]. Consider for simplicity a MIMO radar exploiting a sparse linear array (SLA) which is formed from a uniform linear array (ULA) with half-wavelength interelement spacing by selecting a subset of ULA antennas but maintaining the same aperture [13]. In this context, the primary interest of sparse array design and processing lies in selecting the surviving array elements and carrying out direction finding with the virtual sparse array. Interpolation and extrapolation techniques are widely adopted in automotive radars to fill the holes in the synthesized SLA [3], [16].

In this paper, we propose a high-fidelity automotive MIMO radar sensing technique that exploits a sparse two-dimensional (2-D) MIMO array and provides four-dimensional (4D) point clouds at a much lower cost than LiDAR and with higher robustness to weather conditions. Each transmit antenna transmits the same random sparse step-frequency waveform (RSSFW) to synthesize a large effective bandwidth for high-resolution range estimation, while keeping a low sampling rate. In addition, sparse spectrum utilization makes it robust to multiuser interference from other automotive radars. The waveform orthogonality between MIMO transmit antennas is achieved through Doppler division multiplexing (DDM). At the receiver, targets are first distinguished in the joint range-Doppler domain, and the large virtual sparse array with hundreds of elements synthesized from the 2D MIMO radar is used to provide high-resolution image in both azimuth and elevation. We further use matrix completion techniques [17]–[21] to complete the corresponding virtual uniform rectangular arrays (URAs), and carry...
out high-resolution direction finding.

II. SPARSE STEP-FREQUENCY AUTOMOTIVE RADAR

In this section, we address the problem of high-resolution target range estimation using RSSFW signals with a small number of carrier frequencies. A transmit antenna transmits a sequence of \( N \) pulses whose carrier frequencies \( f_n \in [f_c, f_c + B], \) \( n = 1, 2, \ldots, N, \) are randomly chosen from the set \( \mathcal{M} = \{ f_n = f_c + h_n \Delta f, h_n \in \{1, 2, \ldots, P\} \} \) with \( P = \lfloor B/\Delta f \rfloor \) equally spaced subcarriers, where \( \Delta f \) is the frequency step size and \( \lfloor \cdot \rfloor \) denotes the floor function. The unambiguous scope of high range resolution profiles (HRRP) and range resolution are respectively given by \( R_{\text{high}} = c/(2\Delta f) \) and \( \Delta R = c/(2P\Delta f) = R_{\text{high}}/P. \) The duration of each pulse is \( T_p. \) After a burst of \( N \) pulses are transmitted, the radar is switched to a receive mode. The total time duration of a burst cycle consisting of both transmit and receive modes is \( T. \) The maximum unambiguous range is \( R_{\text{max}} = cT/2. \) One CPI consists of \( M \) burst cycles. The \( n \)-th transmit pulse during the \( n \)-th burst cycle is expressed as

\[
s(m, n, t) = \frac{1}{\sqrt{P}} \text{rect} \left( \frac{t - nT_p - mT_p}{T_p} \right) e^{j2\pi(t-nT_p-mT)t}, \tag{1}
\]

where \( t \) is the fast time, and

\[
\text{rect} \left( \frac{t - \tau}{T_p} \right) = \begin{cases} 1, & \tau \leq t \leq \tau + T_p, \\ 0, & \text{otherwise}. \end{cases} \tag{2}
\]

Each pulse has unit energy, i.e., \( \int_0^{T_p} |s(m, n, t)|^2 dt = 1. \)

Consider \( K \) target points in the far field, the \( k \)-th target has range \( r_k, \) radial velocity \( v_k, \) and complex reflection coefficient \( \beta_k. \) The received signal of the \( n \)-th pulse at the \( m \)-th slow time corresponding to the \( k \)-th target is

\[
y_k(m, n, t) = \beta_k s(m, n, t - 2r_k(t)/c), \tag{3}
\]

where \( r_k(t) = r_k(0) + v_k t \) and \( c \) is the speed of light. After demodulation, the \( n \)-th echo is sampled at the rate of \( 1/T_p \) at fast time \( t_s(m, n) = mT + nT_p \) with \( n = 1, 2, \ldots, N, \) giving one sample per frequency step, expressed as

\[
y_k(m, n) = \beta_k e^{-j\frac{2\pi}{T_p} f_n[r_k(0) + (mT + nT_p)v_k]}. \tag{4}
\]

The sampled received signal for the \( n \)-th pulse is the superposition of the echoes from all \( K \) targets, i.e.,

\[
y(m, n) = \sum_{k=1}^{K} y_k(m, n) = \sum_{k=1}^{K} \beta_k e^{-j\frac{2\pi}{T_p} f_n[r_k(0) + (mT + nT_p)v_k]}
\]

\[
= \sum_{k=1}^{K} \beta_k e^{-j\frac{2\pi}{T_p} f_n[r_k(0) + (mT + nT_p)v_k]}
\]

\[
= \sum_{k=1}^{K} \beta_k e^{-j\frac{2\pi}{T_p} f_n[r_k(0) + (mT + nT_p)v_k]}
\]

\[
\times e^{-j\frac{2\pi}{T_p} f_n[r_k(0) + (mT + nT_p)v_k]}. \tag{5}
\]

where \( \gamma_k = \beta_k e^{-j\frac{2\pi}{T_p} f_n[r_k(0)]}. \)

We make the following assumptions:

A1) The unambiguous scope of high range resolution profiles, defined as \( c/(2\Delta f), \) is larger than the scope of a range bin \( cT_p/2. \) This yields that \( \Delta f < 1/T_p. \)

A2) The range migration is negligible during one CPI, i.e., \( v_k T_p < cT_p/2. \)

A3) Considering a typical vehicle speed, it is reasonable to assume \( \xi_{m,n} = 2(2n\Delta f T_p + m\Delta f(T + 2f_c T_p))v_k/c < 1/P \) for \( m = 0, \ldots, M - 1 \) and \( n = 0, \ldots, P - 1. \)

A4) The Doppler shift is considered constant in one burst cycle \( T \) because of the short duration of the burst pulses.

Range estimation can be achieved by applying inverse discrete Fourier transform (IDFT) to the fast-time samples. The range resolution is determined by the frequency bandwidth. Traditional step-frequency radar systems require \( N = P \) pulses to achieve a range resolution of \( R_{\text{high}}/P. \) For the proposed sparse step-frequency approach, we use \( N < P \) pulses and still achieve the same range resolution of \( R_{\text{high}}/P. \) For each range bin \( l, \) the velocity estimation is obtained by applying discrete Fourier transform (DFT) to the obtained range spectra. P-point IDFT for range estimation and M-point DFT for Doppler estimation provide \( 10\log_{10}(PM) \) dB SNR enhancement [4]. This SNR enhancement acts as a processing gain which significantly benefits the subsequent angle estimation.

We extend the RSSFW radar to a MIMO setting through DDM [4] by applying phase codes along the slow time so that the Doppler DFT of the interference can be distributed into the entire Doppler spectrum and treat it as pseudo noise with a low power spectrum.

III. TWO-DIMENSIONAL SPARSE ARRAYS DESIGN

Depending on the performance and cost requirements, automotive radar can use one or multiple transceivers to synthesize an SLA for direction finding. The utilization of sparse arrays reduces not only the hardware cost but also mutual coupling effects. The array response at a particular time instance consisting of data obtained at all the \( M, M_r \) virtual receivers and corresponding to the same range-Doppler bin is defined as an array snapshot. To mitigate the side lobes introduced by the sparse arrays, we utilize the matrix completion technology to interpolate/extrapolate the holes in the sparse arrays. Matrix completion effectively improves the SNR of array response as the array holes are fully recovered.

The success of applying matrix completion in sparse arrays relies on the following two facts:

F1) The number of targets in the same range-Doppler bin that need angle estimation is small and sparsely present so the targets are first separated in the joint range-Doppler domain.

F2) The SNR in the array snapshot is high because of the processing gain in both range and Doppler domains. The high SNR help reduce the matrix completion error and improve the accuracy of angle estimation.

To enable drive-over and drive-under functions, the elevation angles of targets must be precisely measured with high-resolution angle discrimination capability. Fig. 1 shows a MIMO radar with 12 transmit antennas and 16 receive antennas that are obtained by cascading 4 automotive radar transceivers consisting of 3 transmit antennas and 4 receive antennas. The transmit and receive antennas are randomly deployed on a half-wavelength grid in an area of \([0, 100] \times [0, 120] \times [0, 77] \) to synthesize a MIMO 2D virtual sparse array of 196 elements. The 2D physical array corresponds to a form factor of about \( 20 \times 24 \) cm when the carrier frequency is \( f_c = 77 \text{ GHz}. \) The dimension of the rendered 2D virtual sparse array is \( D_x \times D_y = 183(\lambda/2) \times 194(\lambda/2), \) which can be viewed as a spatial sub-Nyquist sampling of a URA of the same
dimension with half-wavelength spacing in both horizontal and vertical directions. The azimuth and elevation angular resolutions are expressed as $\Delta \theta_{AZ} = 2\arcsin \left( \frac{1.4 \lambda}{\pi D_p} \right) \approx 0.53^\circ$ and $\Delta \theta_{EL} = 2\arcsin \left( \frac{1.4 \lambda}{\pi D_p} \right) \approx 0.56^\circ$, respectively [22]. The angular resolution of imaging radar in this example is comparable to the Velodyne LiDAR HDL-32E whose horizontal and vertical resolutions are 0.16$^\circ$ and 1.33$^\circ$, respectively [23].

Consider a general case of an $M_1 \times M_2$ URA with half-wavelength spacing, shown in Fig. 2, where the URA is located on the $x$-$y$ plane. Assume the $k$-th point target with azimuth angle $\theta_k$ and elevation angle $\phi_k$. Let $\chi_k$ denote the angle between the $k$-th target and the $x$ axis, and $\phi_k$ denote the angle between the $k$-th target and the $y$ axis. Then, it holds that $\cos (\chi_k) = \sin (\phi_k) \cos (\theta_k)$, $\cos (\phi_k) = \sin (\phi_k) \sin (\theta_k)$. Therefore, the azimuth angle $\theta_k$ and elevation angle $\phi_k$ can be uniquely estimated from $\chi_k$ and $\phi_k$ as

$$\theta_k = \arctan \left( \frac{\cos (\phi_k)}{\cos (\chi_k)} \right),$$

$$\phi_k = \arcsin \left( \sqrt{\cos^2 (\chi_k) + \cos^2 (\phi_k)} \right).$$

The $(m_1, m_2)$-th element of the URA array on the $x$-$y$ plane response with respect to the $K$ targets with angle $\chi_k$ to the $x$-axis and angle $\phi_k$ to the $y$-axis, $k = 1, ..., K$, can be written as

$$x_{m_1, m_2} = \sum_{k=1}^{K} \beta_k e^{j \pi ((m_1-1) \sin (\chi_k) + (m_2-1) \sin (\phi_k))}$$

for $1 \leq m_1 \leq M_1$ and $1 \leq m_2 \leq M_2$. Let $\mathbf{M} = [x_{m_1, m_2}]_{0 \leq m_1 \leq M_1, 0 \leq m_2 \leq M_2}$ be the data matrix with entries as the URA array response defined in (8).

We can construct an $N_1 \times (M_1 - N_1 + 1)$ block Hankel matrix as

$$\mathbf{Y}_E = \begin{bmatrix} Y_0 & Y_1 & \cdots & Y_{M_1-N_1} \\ Y_1 & Y_2 & \cdots & Y_{M_1-N_1+1} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{N_1-1} & Y_{N_1} & \cdots & Y_{M_1-1} \end{bmatrix},$$

where

$$\mathbf{Y}_m = \begin{bmatrix} x_{m,0} & x_{m,1} & \cdots & x_{m, M_2-L} \\ x_{m,1} & x_{m,2} & \cdots & x_{m, M_2-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m, L-1} & x_{m, L} & \cdots & x_{m, M_2-1} \end{bmatrix}$$

is an $L \times (M_2 - L + 1)$ Hankel matrix whose elements are defined in (8). It can be verified that the rank of matrix $\mathbf{Y}_E$ is $K$ if $N_1 \geq K$ and $L \geq K$ [24].

By designing the locations of the transmit and receive antennas, we aim to synthesize a sparse 2D array, which can be viewed as spatial subsampling of the block Hankel matrix $\mathbf{Y}_E$ corresponding to a URA. Under certain conditions, the missing elements can be fully recovered by solving a relaxed nuclear norm optimization problem conditioned on the array response of sparse arrays, i.e.,

$$\min \| \mathbf{X}_E \|_* \quad \text{s.t.} \quad \mathcal{P}_\Omega (\mathbf{X}) = \mathcal{P}_\Omega (\mathbf{M})$$

where $\| \cdot \|_*$ denotes the nuclear norm of a matrix, $\mathcal{P}_\Omega (\mathbf{X})$ is the sampling operator with $\Omega$ denoting the observation set consisting of the location of 2D sparse virtual array elements, and $\mathbf{X}_E$ is the block Hankel matrix constructed from matrix $\mathbf{X}$ following equations (9) and (10). In the noisy observation scenario, $\mathbf{M}$ is replaced by $\mathbf{M}^e = [x_{m_1, m_2}]_{0 \leq m_1 \leq M_1, 0 \leq m_2 \leq M_2}$ with $x_{m_1, m_2}^e = x_{m_1, m_2} + n_{m_1, m_2}$, where $x_{m_1, m_2}$ denotes the observed signal and $n_{m_1, m_2}$ is the noise term. We assume the noise is bounded, i.e., $\| \mathcal{P}_\Omega (\mathbf{E}) \|_F \leq \delta$. The block Hankel matrix completion problem of the noisy signal is formulated as

$$\min \| \mathbf{X}_E \|_* \quad \text{s.t.} \quad \| \mathcal{P}_\Omega (\mathbf{X} - \mathbf{M}^e) \|_F \leq \delta.$$ 

Once the block Hankel matrix is completed, higher-resolution direction finding can be achieved with subspace based methods, such as MUSIC [25] and ESPRIT [26], and compressive sensing based methods [27]–[30].
IV. NUMERICAL RESULTS

In one burst cycle, \( N = 300 \) pulses are transmitted. The start carrier frequency is \( f_c = 77 \) GHz, and the effective bandwidth is set to \( B = 200 \) MHz, corresponding to range resolution of \( \Delta R = 0.75 \) m. The pulse duration is \( T_p = 25 \) ns and the step frequency is \( \Delta f = 0.5 \) MHz. The maximum unambiguous detectable range is \( R_u = 300 \) m. The burst cycle repetition interval is \( T = 25 \) \( \mu \)s. The maximum unambiguous detectable velocity is \( v_{\text{max}} = \lambda/(4T) \approx 38.96 \) m/s. To measure the target velocity, \( M = 300 \) burst cycles are carried out with a dwell time of \( MT = 7.5 \) ms, rendering a velocity resolution of \( \Delta v = \lambda/(2MT) = 0.26 \) m/s. To achieve waveform orthogonality among transmit antennas, a Chu sequence [31] of length \( M = 307 \) was generated and then truncated into length \( M = 300 \) for phase coding in slow time. The SNR of the demodulated echo signals at receiver is set to 10 dB.

IV-A. Range-Doppler Spectrum under RSSFW

We consider two targets with equal radar cross section at the same range of \( R = 100 \) m and with the same velocity of \( v = -10 \) m/s. They have different angles represented by \((\chi_1, \varphi_1) = (-20^\circ, 5^\circ), (\chi_2, \varphi_2) = (20^\circ, 10^\circ)\), respectively. To demonstrate waveform orthogonality of RSSFW through DDM, we consider a simple two transmit antenna scenario. Fig. 3 shows the range-Doppler spectrum of the two targets with low sidelobe levels after applying range weighting. There is a flat Doppler ridge which is the residual from the other transmit antennas after demodulation in slow time. The input SNR of the array response after range-Doppler processing is set to 20 dB. We construct a block Hankel matrix \( \mathbf{Y}_R \) of dimension \( 9,009 \times 8,928 \) using one array snapshot. Only \( 0.78\% \) of the Hankel matrix entries are non-zero. The block Hankel matrix is completed via the singular value thresholding (SVT) algorithm [32] to obtain all entries of the corresponding full URA. In this simulation setting, matrix completion contributes around \( 10 \log 10 ((35,502/196)) \approx 22.5 \) dB SNR gain for array processing.

In this simulation setting, matrix completion contributes around \( 10 \log 10 ((35,502/196)) \approx 22.5 \) dB SNR gain for array processing.

Figs. 4 and 5 plot the azimuth-elevation spectra of the two targets under the 2D sparse array and the completed full URA, respectively. It is found that both sparse array and completed URA generate two peaks corresponding to the correct target azimuth and elevation angles. However, the high sidelobes of the sparse array over the entire azimuth and elevation FOVs are effectively mitigated in the completed URA.

IV-B. Two-Dimensional Sparse Array Completion

Now we consider the same two-target scene but use the 2D physical array shown in Fig. 1 for joint high-resolution azimuth and elevation angle estimation. The dimension of the 2D sparse array shown in Fig. 1 is \( D_x \times D_z = 183(\lambda/2) \times 194(\lambda/2) \), which requires \( 35,502 \) elements if a URA of the same dimension is used. The virtual sparse array corresponds to only 0.54\% of the total elements of the URA.

The input SNR of the array response after range-Doppler processing is set to 20 dB. We construct a block Hankel matrix \( \mathbf{Y}_R \) of dimension \( 9,009 \times 8,928 \) using one array snapshot. Only 0.78\% of the Hankel matrix entries are non-zero. The block Hankel matrix is completed via the singular value thresholding (SVT) algorithm [32] to obtain all entries of the corresponding full URA. In this simulation setting, matrix completion contributes around \( 10 \log 10 ((35,502/196)) \approx 22.5 \) dB SNR gain for array processing.

Figs. 4 and 5 plot the azimuth-elevation spectra of the two targets under the 2D sparse array and the completed full URA, respectively. It is found that both sparse array and completed URA generate two peaks corresponding to the correct target azimuth and elevation angles. However, the high sidelobes of the sparse array over the entire azimuth and elevation FOVs are effectively mitigated in the completed URA.

V. CONCLUSIONS

In this paper, we developed an 4D automotive MIMO radar imaging system exploiting a joint thinned frequency spectrum and sparse arrays to synthesize a large effective bandwidth for high-resolution range-Doppler profiles, as well as 2D sparse arrays with hundreds of virtual elements for angle estimation in both azimuth and elevation. The missing elements in the sparse arrays are effectively interpolated using matrix completion, rendering low-sidelobe angle estimation.
VI. REFERENCES


