

Cooperative Robotics Visible Light Positioning: An Intelligent Compressed Sensing and GAN-Enabled Framework

Sicong Liu, *Senior Member, IEEE*, Xianyao Wang, Jian Song, *Fellow, IEEE*, and Zhu Han, *Fellow, IEEE*

Abstract—This paper presents a compressed sensing (CS) based framework for visible light positioning (VLP), designed to achieve simultaneous and precise localization of multiple intelligent robots within an indoor factory. The framework leverages light-emitting diodes (LEDs) originally intended for illumination purposes as anchors, repurposing them for the localization of robots equipped with photodetectors. By predividing the plane encompassing the robot positions into a grid, with the number of robots being notably fewer than the grid points, the inherent sparsity of the arrangement is harnessed. To construct an effective sparse measurement model, a sequence of aggregation, autocorrelation, and cross-correlation operations are employed to the signals. Consequently, the complex task of localizing multiple targets is reformulated into a sparse recovery problem, amenable to resolution through CS-based algorithms. Notably, the localization precision is augmented by inter-target cooperation among the robots, and inter-anchor cooperation among distinct LEDs. Furthermore, to fortify the robustness of localization, a generative adversarial network (GAN) is introduced into the proposed localization framework. The simulation results affirm that the proposed framework can successfully achieve centimeter-level accuracy for simultaneous localization of multiple targets.

Index Terms—Robotics sensing, visible light positioning, multi-target localization, compressed sensing, cooperative localization.

I. INTRODUCTION

In recent years, driven by the wave of integrated sensing and communication (ISAC), industrial Internet of Things (IIoT) technology has developed rapidly [1]. This evolution is driven by the continuous enhancement and diversification of sensors, controllers, communication systems, signal processing,

This work is supported in part by the Natural Science Foundation of Fujian Province of China (No. 2023J01001), the Open Research Fund of National Mobile Communications Research Laboratory, Southeast University (No. 2023D10), the Basic and Applied Basic Research Foundation of Guangdong Province, the National Natural Science Foundation of China (No. 62077040), the Science and Technology Key Project of Fujian Province, China (No. 2021HZ021004), and the Science and Technology Key Project of Xiamen (No. 3502Z20221027), and in part by NSF CNS-2107216, CNS-2128368, CMMI-2222810, ECCS-2302469, US Department of Transportation, Toyota and Amazon. (*Corresponding Author: Sicong Liu*)

Sicong Liu and Xianyao Wang are with the Department of Information and Communication Engineering, School of Informatics, Xiamen University, Xiamen 361005, China. Sicong Liu is also with National Mobile Communications Research Laboratory, Southeast University, China, and also with Shenzhen Research Institute of Xiamen University, Shenzhen 518057, China (E-mail: liusc@xmu.edu.cn).

Jian Song is with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China, the Beijing National Research Center for Information Science and Technology (BNRist), Beijing 100084, China, and also with the Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China (E-mail: jsong@tsinghua.edu.cn).

Zhu Han is with the Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004 USA, and also with the Department of Computer Science and Engineering, Kyung Hee University, Seoul 446-701, South Korea (E-mail: hanzhu22@gmail.com).

and related technologies, leading to a profound reshaping of the traditional industrial production landscape. Notably, this transformation is evidenced by the pervasive integration of cutting-edge equipment such as automated guided vehicles (AGVs) and automatic robots across various stages of industrial production [2], [3]. For instance, prominent logistics companies have deployed a multitude of logistics sorting robots capable of meticulously scanning express shipment information. These robots adeptly direct packages to their respective destinations, effecting unified loading and distribution for different endpoints. Embracing the smart factory paradigm, contemporary industrial production embodies qualities of automation, profound integration, robust real-time capabilities, and heightened security within an intricately interconnected information ecosystem [4], driving steady progression towards intelligent IIoT that underpins modern production processes.

In the expansive domain of IIoT technology, a multitude of fields are encompassed, each presenting a spectrum of key technical challenges. A pivotal facet of this landscape involves the imperative for achieving high-precision real-time positioning of equipment within specific operational segments. Notably, the precision demands inherent to industrial production underscore the need for elevated levels of positioning accuracy [5]. The efficacy of the system hinges on its capability to concurrently localize multiple targets, ensuring efficient and streamlined operations. Meanwhile, the multifaceted nature of factory environments should be considered, wherein an assemblage of signal-transmitting and signal-receiving equipment is concentrated, and there exists a substantial backdrop of ambient noise and electromagnetic interference, necessitating robust measures to ensure resilience to interference [6].

Generally, a prevalent indoor positioning technique involves the use of location fingerprinting based on received signal strength (RSS) [7]. This method functions by partitioning the space into grids and estimating the target position by comparing the actual measured RSS with stored position fingerprints in a database. However, this approach commonly entails substantial computational costs [8]. Moreover, the accuracy of wireless indoor positioning typically reaches the decimeter level, and the performance will be further degraded in more complex environments due to the vulnerability to multipath fading [9].

In recent times, visible light positioning (VLP) technology, harnessing ubiquitous light-emitting diodes (LEDs) for signal transmission, has experienced rapid growth. This approach boasts key benefits, including heightened efficiency, energy conservation, economical deployment, robust anti-interference capabilities, and exceptional precision [10]. Notably, VLP stands out due to its divergence from wireless channels,

i.e., its primary concentration in line-of-sight (LoS) links empowers it to effectively counteract multipath interference [11]. Classic VLP techniques rely on metrics such as RSS, time-of-arrival (TOA) [12], and angle-of-arrival (AOA) [13], etc. These methodologies deliver centimeter-level accuracy in single-target localization. However, a comprehensive high-precision, multi-target simultaneous VLP system remains an ongoing pursuit.

In the context of a straightforward indoor multi-target scenario, the actual count and spatial extent of these targets tend to remain manageable. Consequently, the positions of these multiple targets exhibit a sparse distribution in comparison to the entirety of the indoor space. This rationale facilitates the introduction of the compressed sensing (CS) theory. This integration serves to reframe the challenge of indoor multi-target localization into a sparse recovery problem amenable to effective resolution through classical algorithms like basis pursuit (BP) and orthogonal matching pursuit (OMP) [14]. However, in the context of certain intricate or specialized problems, traditional CS algorithms such as BP and OMP might fall short of ensuring high precision [15]. To address these limitations, the application of generative adversarial network (GAN) presents itself as a promising avenue. Its capacity to discern intricate mappings within data fosters its integration into the CS framework, offering a solution to sparse recovery quandaries [16], [17], which holds potential for enhancing the performance of CS-based VLP schemes, particularly in intricate conditions.

Consequently, to address the limitations of existing positioning methods and achieve simultaneous localization of multiple intelligent robot targets, a CS-based multi-target cooperative VLP framework is devised. Specifically, we implement a downlink VLP utilizing the existing lighting system. By solving a sparse recovery problem well-suited for CS-based algorithms, we accurately determine the positions of all intelligent robots. Collaboration among robots, i.e., inter-target cooperation, is pivotal for constructing this CS framework. Moreover, to enhance positioning accuracy in intricate indoor settings, we leverage correlations between visible light signals from distinct LEDs, thus fostering inter-anchor cooperation. Furthermore, we enhance this scheme by integrating a GAN into our proposed framework, which not only reduces positioning errors but also bolsters noise resistance in complex environments. In summary, our work presents several key contributions as follows:

- **CS-based Multi-Target VLP Framework:** We introduce a multi-target VLP framework founded on CS, leveraging preexisting indoor lighting systems, which facilitates concurrent high-precision positioning of multiple targets.¹

¹Part of this work has been accepted for publication in IEEE International Conference on Communications (IEEE ICC) 2023 [18]. Compared with the short conference version, this article has extensively extended the technical content, theoretical analysis and experimental results. An enhanced CS-GAN enabled localization method is further proposed to deal with the multi-target VLP task, notably enhancing the precision of sparse recovery and the positioning accuracy in severe conditions. Theoretical analysis of the positioning performance bound is derived. More extensive and thorough simulations have been conducted and reported and discussed on to evaluate the proposed schemes.

- **Cooperation among Targets and Anchors:** By capitalizing on the interplay between the robot targets and the LED anchors, the precision of multi-target positioning outcomes is elevated.
- **Enhanced CS-GAN Enabled Localization:** GAN is incorporated into the multi-target VLP framework, notably enhancing the precision of sparse recovery and the positioning accuracy. Remarkably, even when operating under low signal-to-noise ratio (SNR), the fortified CS-GAN approach maintains a high level of accuracy in estimating multiple robot positions.

The rest of this article is structured as follows: Section II investigates related work. Section III describes a system model for multi-objective visible light localization in an indoor factory. Section IV introduces the CS-based multi-target cooperative VLP framework we proposed. Section V proposes to apply GAN to the proposed CS-based multi-target cooperative VLP framework to further improve positioning performance. Section VI provides a theoretical analysis and performance evaluation of the proposed method. The simulation results are given in Section VII, and the final conclusions are given in Section VIII.

Notation. Matrices and vectors are denoted by boldface letters; $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ denote the complex-conjugate operation, the transpose operation and the conjugate transpose operation, respectively; $\|\cdot\|_r$ represents the ℓ_r -norm operation.

II. RELATED WORKS

IoT is a creative and promising paradigm for various applications of robotics. Some works have looked ahead to the changes that will be brought about in practical applications. A study on the reconfigurability of robots in various industrial automation processes proposed a region based environment prediction method, which effectively improved the detection accuracy and speed of robots in identifying different task targets [19]. In order to improve the positioning accuracy of mobile robots in complex indoor environments, Xin *et al* proposed a multi-mobile robot collaborative positioning system based on ultra-wideband sensor and GPU hardware acceleration to reduce the influence of non-line-of-sight error [20]. A novel charging robot called DeltaCharger has been designed for 3D positioning of electrodes [21], which can achieve an accuracy of over 90% via the convolutional neural network.

For the problem of indoor positioning, radio frequency based wireless positioning schemes have been extensively studied. In order to improve the accuracy and versatility of wireless positioning for IoT terminals, Yu *et al* proposed a highly integrated hybrid wireless positioning system [22]. Gnlta *et al* used channel state information for fingerprint identification and proposed a wireless positioning framework [23]. Li *et al* proposed an unsupervised wireless localization method based on deep reinforcement learning without knowing the labels of actual positions [24].

In recent years, with the development of visible light communications (VLC), many VLP methods have been proposed to achieve higher precision indoor positioning. In order to

improve the robustness of VLP system to different models, Keskin *et al* derived the Cramr-Rao Lower bound (CRLB) and designed direct and two-step estimators for synchronous and asynchronous positioning [25]. In order to combine positioning and communication, Wei *et al* formulated a visible light integrated positioning and communication framework to improve the spectrum efficiency and the overall system performance [26]. Zhou *et al* comprehensively investigated the impact of SNR, non LoS propagation and other factors on the performance of RSS-based VLP, and derived the performance limits for it [27].

Due to the sparse nature of the target location in comparison to the entirety of the indoor space, the CS technology has been utilized in positioning. For indoor wireless positioning, N. Garcia *et al* proposed a method called direct source localization that enables robust localization with decimeter-level accuracy in massive MIMO systems [28]. Jamali-Rad *et al* exploited the latent information in received signal correlations by introducing a new fingerprint paradigm and then used CS-based fingerprinting methods to solve the problem of locating multiple sources in multipath environments [29]. To improve the efficiency of radio map measuring and overcome the heterogeneity due to different mobile devices, Gong *et al* proposed a CS-based fingerprint localization method [30]. There are also some works that applied CS in VLP. K. Gligori *et al* transferred the LED signal separation problem into an equivalent CS model and proposed an indoor VLP method using the proximity method [31]. R. Zhang *et al* proposed a reverse visible light multi-target localization method based on sparse matrix reconstruction [32], while this model is different from the downlink VLC scenario using lighting-purpose LEDs commonly deployed in practice.

Recently, the emerging deep learning technology is being considered with CS, in order to learn the sparse features of the signals and improve the sparse recovery accuracy. Wei *et al* proposed a privacy-aware sensing and transmission scheme for the internet of medical things, and devised a sparse-learning-based encryption and recovery method to protect the privacy and reduce power consumption [33]. In order to improve the performance of wireless vehicular communication system based on orthogonal frequency division multiplexing (OFDM), Liu *et al* proposed a sparse machine learning scheme to eliminate the impulse noise [34]. A deep learning based scheme was proposed for a hybrid analog-digital massive MIMO system, which can efficiently implement the uplink channel estimation and reduce the computational complexity [35].

III. SYSTEM MODEL AND PROBLEM FORMULATION

A. Channel Model of Visible Light Propagation

In the VLP system, lighting LEDs are generally used directly as the light signal source. The irradiation intensity R_o follows the Lambertian radiation pattern [36], which is expressed as

$$R_o(\alpha) = \frac{m+1}{2\pi} \cos^m(\alpha), -\frac{\pi}{2} \leq \alpha \leq \frac{\pi}{2}, \quad (1)$$

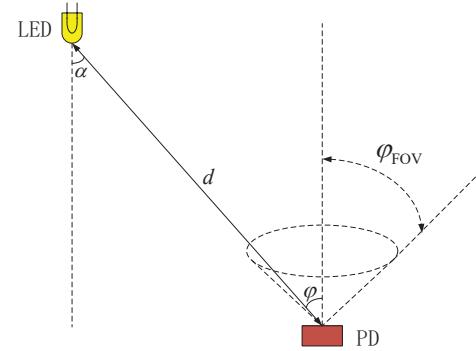


Fig. 1: Channel model of visible light signal propagation in an indoor VLP scenario.

where α is the irradiation angle of the LED, and m is the Lambertian radiation ordinal given by

$$m = -\ln 2 / \ln (\cos \alpha_{1/2}), \quad (2)$$

where $\alpha_{1/2}$ denotes the half-power angle [37]. The robotics terminals are equipped with photodetectors (PDs) to detect the visible light signals. As shown in Fig. 1, considering the LoS link, the channel gain of visible light propagation is expressed as

$$h = \begin{cases} \frac{1}{d^2} R_o(\alpha) A_{\text{eff}}(\varphi), & 0 \leq \varphi \leq \varphi_{\text{FOV}} \\ 0, & \text{else} \end{cases}, \quad (3)$$

where d is the distance between the LED and the PD; φ is the angle of incidence at the PD; φ_{FOV} is the field of view (FOV) of the PD; $A_{\text{eff}}(\varphi)$ is the effective detection area of the PD, which is given by

$$A_{\text{eff}}(\varphi) = A_{\text{det}} G_{\text{filter}} G_{\text{conc}} \cos \varphi, \quad (4)$$

where A_{det} is the physical detection area of the detector; G_{filter} and G_{conc} represent the gain of the optical filter and the gain of the optical concentrator, respectively [38].

B. System Model of Indoor Multi-Target Visible Light Positioning

To achieve a concise and clear representation, we simplify the indoor VLP positioning scenario involving multiple intelligent robots as the targets to be located, as depicted in Fig. 2. In this arrangement, several ceiling-mounted lighting LEDs emit visible light signals. The downlink VLC transmission is based on the intensity-modulation/direct-detection (IM/DD) technique. A multi-carrier modulation scheme, such as DC biased optical OFDM (DCO-OFDM), i.e., classical OFDM with a DC bias ensuring nonnegative value, can be adopted for the VLC transmission of these LEDs, where different OFDM subcarriers can be allocated to different LEDs to ensure multiple access with high spectral efficiency. The information of unique identification and accurate location of each LED is carried in the downlink VLC signal. Intelligent robots equipped with PDs traverse a two-dimensional plane suspended above the floor by a distance of d_h . Notably, this

plane is systematically partitioned into N grid points.² Each of these grid points signifies a potential position for an intelligent robot, corresponding to moments of operation.

To facilitate the calculation of the channel gain, a coordinate system covering the entire room is established, so that the position of each grid point can be represented in three-dimensional coordinates. Let M denote the number of LEDs, and thus the coordinates of the LEDs can be expressed as

$$(x_i^{(\text{tx})}, y_i^{(\text{tx})}, z_i^{(\text{tx})}), i = 1, \dots, M. \quad (5)$$

Let P_j denote the j -th grid point, and thus the coordinates of all the N grid points can be expressed as

$$(x_j^{(\text{rx})}, y_j^{(\text{rx})}, z_j^{(\text{rx})}), j = 1, \dots, N. \quad (6)$$

Then, the distance d_{ij} from the i -th LED to the grid point P_j can be represented by

$$d_{ij} = \left((x_i^{(\text{tx})} - x_j^{(\text{rx})})^2 + (y_i^{(\text{tx})} - y_j^{(\text{rx})})^2 + (z_i^{(\text{tx})} - z_j^{(\text{rx})})^2 \right)^{1/2}. \quad (7)$$

Substituting (7) into (3), the channel gain h_{ij} from the i -th LED to the grid point P_j can be obtained. Let x_i represent the pilot signal sent by the i -th LED, which is utilized for localization purpose. Suppose an intelligent robot is located at the grid point P_j . Then, the pilot signal received by the robot at P_j and sent from the i -th LED can be denoted as

$$y_{ij} = h_{ij}x_i + \omega_{ij}, \quad (8)$$

where ω_{ij} represents the background Additive White Gaussian Noise (AWGN) in this link.

As illustrated in Fig. 2, the positions of the intelligent robots are inherently random and undisclosed. The primary objective of the multi-target positioning endeavor is to ascertain the specific grid points corresponding to the positions of intelligent robots, which is accomplished by analyzing the light signals received by the various intelligent robots from distinct LEDs. In practice, the quantity of intelligent robots is relatively modest, notably fewer than the total count of grid points, where the distribution of the robots can be arbitrarily random without a prior constraint. This sparse distribution of robots aligns well with the principles of CS, by which we can effectively fulfill the multi-target positioning task.

IV. MULTI-TARGET COOPERATIVE VISIBLE LIGHT POSITIONING FRAMEWORK VIA COMPRESSED SENSING

In this section, we first introduce the details of the formulation of the CS-based multi-target VLP framework enabled by inter-target cooperation in Section IV-A. Then, the framework is enhanced by introducing the inter-anchor cooperation among distinct LEDs, which is described in Section IV-B.

A. Compressed Sensing Based Multi-Target Visible Light Positioning with Inter-Target Cooperation

In this sub section, a framework of CS-based multi-target VLP (CSM-VLP) is formulated to realize accurate multi-target

²The value of d_h does not need to be specific, and the model can be applied to different heights of the PDs.

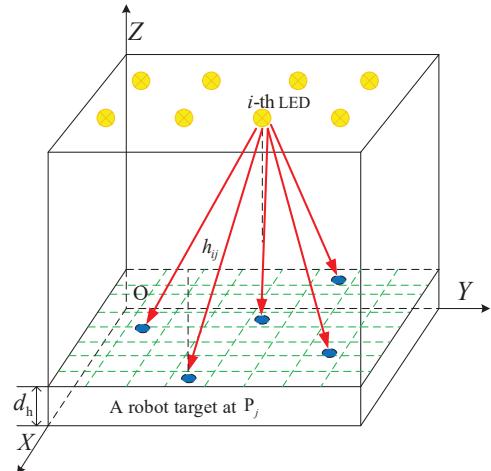


Fig. 2: Geometric layout of an indoor multi-target VLP system: LEDs are utilized as visible light signal transmitters; Multiple intelligent robots equipped with PDs are randomly distributed in the plane that is divided into many grid points.

localization in a typical downlink VLP system. Specifically, assume that the number of intelligent robots is K , which is far smaller than the number of grid points in the area, i.e., $K \ll N$. As described in the previous section, the i -th LED broadcasts the pilot signal x_i to all the robots via the visible light downlink. We use the symbols $\{y_{ik}\}_{k=1}^K$ to denote the pilot signals sent from the i -th LED and received by the K robots. Specifically, y_i denotes the aggregated received signal corresponding to the i -th LED, which is obtained by aggregating the pilot signals received by all the K robots via inter-target cooperation among the robots through, for instance, a wireless link. Therefore, the aggregated received signal $y_i, i = 1, \dots, M$ is given by

$$y_i = \sum_{k=1}^K y_{ik} = \sum_{k=1}^K h_{ik}x_i + \omega_i, \quad (9)$$

where ω_i represents the aggregated background noise. Then, we can formulate an aggregated received signal vector \mathbf{y} by combining all the aggregated received signals $\{y_i\}_{i=1}^M$ corresponding to all the M LEDs, which is given by

$$\mathbf{y} = (y_1, \dots, y_M)^T = \left(\sum_{k=1}^K h_{1k}x_1, \dots, \sum_{k=1}^K h_{Mk}x_M \right)^T + \boldsymbol{\omega}, \quad (10)$$

where $\boldsymbol{\omega}$ is the aggregated background noise vector. Then, we can rewrite (10) into matrix format, leading to a CS-based framework of multi-target VLP as given by

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{bmatrix} = \underbrace{\begin{bmatrix} h_{11}x_1 & \cdots & h_{1N}x_1 \\ h_{21}x_2 & \cdots & h_{2N}x_2 \\ \vdots & \ddots & \vdots \\ h_{M1}x_M & \cdots & h_{MN}x_M \end{bmatrix}}_{\Phi} \boldsymbol{\Theta} + \boldsymbol{\omega}, \quad (11)$$

where $\boldsymbol{\Theta} = [\theta_1, \dots, \theta_j, \dots, \theta_N]^T$ denotes an on-grid localization vector. It can be actually regarded as an indicator vector

with N entries, with each entry valued either one or zero. The value of each entry indicates whether there is a robot located in the corresponding grid point or not. Consequently, most of the entries are zeros, except for the entries corresponding to the grid points the K robots are located. Since the number of the nonzero elements in Θ is much smaller than its length, Θ is a sparse vector. Thus, it can be recovered through sparse recovery methods, including typically CS-based algorithms. In the CS-based measurement model formulated in (11), we can define the observation matrix Φ as

$$\Phi = \begin{bmatrix} h_{11}x_1 & \cdots & h_{1N}x_1 \\ h_{21}x_2 & \cdots & h_{2N}x_2 \\ \vdots & \ddots & \vdots \\ h_{M1}x_M & \cdots & h_{MN}x_M \end{bmatrix}. \quad (12)$$

Next, we can perform autocorrelation on the aggregate received signal vector \mathbf{y} , which yields the aggregated power measurement vector \mathbf{p}_{rx} as given by

$$\begin{aligned} \mathbf{p}_{\text{rx}} &= \mathbb{E}\{\mathbf{y} \odot \mathbf{y}^*\} \\ &= \mathbb{E}\{(\Phi\Theta + \boldsymbol{\omega}) \odot (\Phi\Theta + \boldsymbol{\omega})^*\} \\ &= \mathbb{E}\{\Phi \odot \Phi^*\}\Theta + \mathbb{E}\{\boldsymbol{\omega} \odot \boldsymbol{\omega}^*\} \\ &= \mathbf{J}\Theta + \sigma_n^2 \mathbf{1}_M, \end{aligned} \quad (13)$$

where $\mathbb{E}\{\cdot\}$, \odot , $(\cdot)^*$ represent expectation operator, Hadamard product operator, and complex-conjugate operator, respectively; σ_n^2 denotes noise variance; $\mathbf{1}_M$ is an all-one-valued length- M vector. The purpose of this autocorrelation operation is to turn the CS-based model in (11) to an adaptive refined CS-based power measurement model in (13), which is independent of different pilot signal characteristics. Notably, the observation matrix \mathbf{J} in (13) is given by

$$\mathbf{J} = \begin{bmatrix} |h_{11}|^2 & |h_{12}|^2 & \cdots & |h_{1N}|^2 \\ |h_{21}|^2 & |h_{22}|^2 & \cdots & |h_{2N}|^2 \\ \vdots & \vdots & \ddots & \vdots \\ |h_{M1}|^2 & |h_{M2}|^2 & \cdots & |h_{MN}|^2 \end{bmatrix}. \quad (14)$$

where the elements in \mathbf{J} are related to the channel gains, affected by the deployment of the LEDs and the indoor layout. Thus, the observation matrix \mathbf{J} can be obtained by the channel information collected beforehand in the fingerprint database of the indoor visible light channels.

Now the localization problem is turned into a sparse recovery problem in (13), which is aimed to recover the on-grid target localization sparse vector Θ after obtaining the aggregated power measurement vector \mathbf{p}_{rx} . Then it can be solved by classical sparse recovery methods. After getting the estimate of Θ , the grid points and the corresponding positions of the robot targets can be subsequently obtained.

In the model of the proposed method, the step-size of the grid points has an influence on the performance. Specifically, a denser grid partition makes the quantization error of the location coordinates smaller, leading to higher positioning resolution. Meanwhile, the sparsity of Θ can be better preserved with denser grids, particularly in case of more robots to be located, otherwise the CS algorithm will fail without the sparsity prior. On the other hand, it is not necessarily true that a denser grid

will lead to a better localization performance. According to the CS theory, the condition $M \geq \mu K \log(N/K)$ should be satisfied to reach a satisfactory sparse recovery performance [39]. If the number of grid points N is too large, it requires a larger amount of measurement data M , i.e., the number of LEDs, which might not be satisfied, resulting in a worse performance. In addition, the increase of N will increase the computational complexity. Hence, the number of M and N needs to be appropriately configured according to the practical situations and requirements.

B. Enhanced Compressed Sensing Based Multi-Target Visible Light Positioning with Inter-Anchor Cooperation

As reported in the CSM-VLP scheme, the inter-robot information has been exploited to formulate the measurement vector for the CS model in (13) via autocorrelation. However, additional untapped information remains at our disposal, poised to further fortify the robustness of localization. This encompasses the latent geometric insights of the indoor setting and the channel-specific attributes related to distinct anchors, i.e., LEDs. In order to achieve a better multi-target joint localization performance, we further devise a cooperative CSM-VLP (CoCSM-VLP) scheme, which make use of the inter-anchor cooperation via the cross-correlation between the aggregated received signals corresponding to different LEDs.

Specifically, we calculate the cross-correlation $\mathbb{E}\{y_i y_j^*\}$ between the aggregated received signals given in (11) to obtain the inter-anchor correlation matrix \mathbf{P}_{corr} , which is given by

$$\begin{aligned} \mathbf{P}_{\text{corr}} &= \mathbb{E}\{\mathbf{y}\mathbf{y}^H\} \\ &= \mathbb{E}\{(\Phi\Theta + \boldsymbol{\omega})(\Phi\Theta + \boldsymbol{\omega})^H\} \\ &= \mathbb{E}\{\Phi\Theta^H\Phi^H\} + \mathbb{E}\{\boldsymbol{\omega}\boldsymbol{\omega}^H\}, \end{aligned} \quad (15)$$

where $(\cdot)^H$ represents the conjugate-transpose operator. Then, through vectorizing both sides of (15), the cross-correlation measurement vector \mathbf{p}_{corr} can be derived and thus the CS-based measurement model is formulated as given by

$$\begin{aligned} \mathbf{p}_{\text{corr}} &= \text{vec}(\mathbf{P}_{\text{corr}}) \\ &= \text{vec}\left(\mathbb{E}\{\Phi\Theta^H\Phi^H\}\right) + \text{vec}\left(\mathbb{E}\{\boldsymbol{\omega}\boldsymbol{\omega}^H\}\right) \\ &= \mathbb{E}\{\Phi^* \otimes \Phi\} \text{vec}(\Theta^H) + \text{vec}(\boldsymbol{\omega}\boldsymbol{\omega}^H) \\ &= \mathbb{E}\{\Phi^* \cdot \Phi\}\Theta + \text{vec}(\sigma_n^2 \mathbf{I}_M), \end{aligned} \quad (16)$$

where \otimes represents the Kronecker product operator; \cdot denotes Khatri-Rao product; \mathbf{I}_M denotes a unit matrix of size $M \times M$.

Equation (16) is essentially a series of linear equations. In fact, it is the independent equations therein that directly influence the sparse recovery performance. The linearly correlated rows in the observation matrix in (16) do not contribute to the independent measurements, but may introduce additional noise that degrades the performance. Therefore, one can delete the redundant linearly correlated rows to reduce computational complexity while improving the sparse recovery performance. Since the cross-correlation matrix \mathbf{P}_{corr} in (15) is symmetrical, one can define a selection matrix \mathbf{S} of size $M(M+1)/2 \times M^2$ to select the independent equations of (16), which is the rows corresponding to the M diagonal and $M(M-1)/2$

upper-diagonal elements of \mathbf{P}_{corr} . Then the CS-based cross-correlation measurement model can be formulated as

$$\begin{aligned}\hat{\mathbf{y}} &= \mathbf{S} \operatorname{vec}(\mathbf{P}_{\text{corr}}) \\ &= \mathbf{S} \mathbb{E}\{\Phi^* \cdot \Phi\} \boldsymbol{\theta} + \mathbf{S} \operatorname{vec}(\sigma_n^2 \mathbf{I}_M) \\ &= \boldsymbol{\Psi} \boldsymbol{\theta} + \boldsymbol{\omega}_{\text{vec}},\end{aligned}\quad (17)$$

where $\boldsymbol{\Psi}$ is the observation matrix of size $M(M+1)/2 \times N$ as given by

$$\boldsymbol{\Psi} = \begin{bmatrix} |h_{11}|^2 & |h_{12}|^2 & \cdots & |h_{1N}|^2 \\ h_{11}^* h_{21} & h_{12}^* h_{22} & \cdots & h_{1N}^* h_{2N} \\ |h_{21}|^2 & |h_{22}|^2 & \cdots & |h_{2N}|^2 \\ \vdots & \vdots & \vdots & \vdots \\ h_{11}^* h_{M1} & h_{12}^* h_{M2} & \cdots & h_{1N}^* h_{MN} \\ \vdots & \vdots & \vdots & \vdots \\ |h_{M1}|^2 & |h_{M2}|^2 & \cdots & |h_{MN}|^2 \end{bmatrix}. \quad (18)$$

This observation matrix can be regarded as a fingerprint database of the proposed CoCSM-VLP scheme. It is observed that the elements in $\boldsymbol{\Psi}$ are also related to the channel gains, so $\boldsymbol{\Psi}$ can be similarly obtained by the collected channel information. Note that the number of independent rows in $\boldsymbol{\Psi}$ is much larger than that of the observation matrix \mathbf{J} in (13), so more independent measurements are available to solve for the unknown sparse localization vector $\boldsymbol{\theta}$. Specifically, the amount of available independent linear equations in (17) has increased from M to $M(M+1)/2$ compared with that in (13), bringing about more benefits for the CS-based multi-target localization.

As previously mentioned, it is required that the amount of measurement data should be greater than $\mu K \log(N/K)$ for satisfactory sparse recovery according to the CS theory. Thus, it is much easier for the cross-correlation measurement model in (17) to satisfy this requirement, since the amount of its available measurement data is $M(M+1)/2$, much larger than M for the auto-correlation measurement model in (13). This will improve the accuracy and reliability of localization, especially in severe conditions such as large number of unknown targets (K is large), a denser grid partitioning (N is large), and intensive background noise (the SNR is low). As far as recovering the sparse localization vector $\boldsymbol{\theta}$ in (17) is concerned, it can also be efficiently solved by using classical sparse recovery methods, such as ℓ_1 -norm minimization methods or CS-based greedy algorithms, similarly as done for CSM-VLP.

V. SPARSITY-AWARE GENERATIVE ADVERSARIAL NETWORK ENABLED MULTI-TARGET VISIBLE LIGHT POSITIONING

In the above proposed CS-based multi-target VLP framework, the recovery accuracy of the on-grid robot localization vector $\boldsymbol{\theta}$ directly determines the accuracy of localization. In both the CSM-VLP and CoCSM-VLP schemes, the traditional CS recovery algorithm is utilized to reconstruct $\boldsymbol{\theta}$. However, the performance of these traditional CS algorithms will be limited under some harsh conditions. In recent years, the GAN technology has been introduced into the CS framework because of its strong capability of learning inherent latent mapping from the data to specific high-level features through adversarial training [40], [41]. In order to improve the accuracy

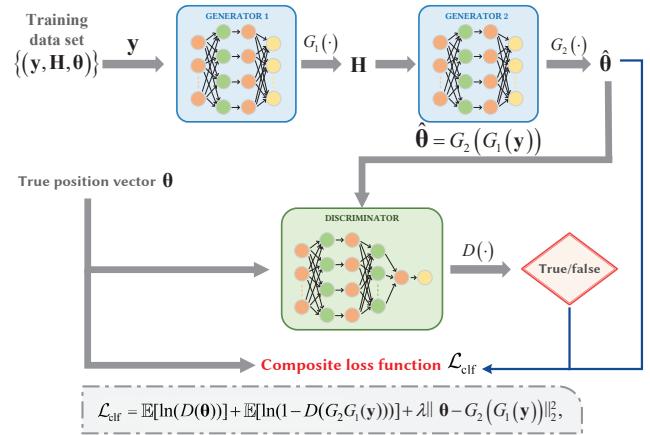


Fig. 3: Schematic diagram of the GAN-based multi-target VLP (GAN-MVLP) scheme.

and robustness of localization, we further integrate GAN into the proposed CS-based localization framework for the specific scenarios of multi-target localization, and propose a GAN-based multi-target VLP (GAN-MVLP) scheme. This method makes use of the capability of the generator network to explore the given sample data, and enables the generator network to directly learn the display mapping relationship from the aggregated received signal vector \mathbf{y} to the sparse localization vector $\boldsymbol{\theta}$ via confrontation training. In addition, in order to improve the recovery accuracy of $\boldsymbol{\theta}$, the proposed GAN-MVLP scheme introduces a regularization term into the traditional loss function to construct a new composite loss function for training the GAN, which can strengthen the ability of learning the sparse features and thus improve the accuracy of multi-target localization.

Specifically, based on the CS-based multi-target VLP framework given in (11), we further build a sparse learning framework based on GAN to recover $\boldsymbol{\theta}$, and its structure is shown in Fig. 3. As shown in Fig. 3, we have made some improvements to the traditional GAN to address the positioning problems to be solved. First, we directly use the aggregated received signal vector \mathbf{y} as the initial input of the generator network, replacing the random variable z in the traditional GAN. Especially, we use two generators to establish the display mapping relationship from \mathbf{y} to $\boldsymbol{\theta}$ step by step. The generator $G_1(\cdot)$ generates the corresponding potential channel matrix \mathbf{H} according to input \mathbf{y} ; Then potential channel matrix \mathbf{H} is fed into the pre-trained generator $G_2(\cdot)$ as the input, which is aimed at generating the corresponding generated sparse localization vector $\hat{\boldsymbol{\theta}}$ containing the location information according to the input channel information.

In addition, we add regularization terms to the traditional loss function to reduce the error between the generated vector and the real vector, and design a composite loss function \mathcal{L}_{clf} , which is defined as

$$\mathcal{L}_{\text{clf}} = \min_{\mathcal{G}} \max_{\mathcal{D}} \mathcal{L}(D, G) + \lambda \|\boldsymbol{\theta} - G_2(G_1(\mathbf{y}))\|_2^2, \quad (19)$$

where

$$\min_{\mathcal{G}} \max_{\mathcal{D}} \mathcal{L}(D, G) = \mathbb{E}[\ln(D(\boldsymbol{\theta}))] + \mathbb{E}[\ln(1 - D(G_2(G_1(\mathbf{y})))])]. \quad (20)$$

Note that $\min_G \max_D \mathcal{L}(D, G)$ is the original countermeasure loss function, which is used to ensure that the vector learned by the two generators has the same statistical distribution as the real positioning vector. $\lambda \|\theta - G_2(G_1(\mathbf{y}))\|_2^2$ is an additional regularization term to reduce the mean square error (MSE) between the generated vector and the real vector, so as to improve the positioning accuracy.

The proposed GAN-MVLP scheme is mainly composed of two stages: training stage and inference stage. The detailed process of the training stage is described in **Algorithm 1**, and summarized are as follows. Firstly, the discriminator network is trained so that it can correctly distinguish the real localization vector θ from the generated localization vector $\hat{\theta}$ to the best extent. Namely, the judgment probability of the real localization vector $D(\theta)$ approaches 1, and the judgment probability of the generated localization vector $D(\hat{\theta})$ approaches 0. The loss function of discriminator network $D(\cdot)$ is given by

$$\max \mathcal{L}_D = \mathbb{E}[\ln(D(\theta))] + \mathbb{E}[\ln(1 - D(G_2(G_1(\mathbf{y}))))]. \quad (21)$$

In the training stage, BP and gradient descent (GD) algorithms are used to update and optimize the learnable parameter Θ_D of $D(\cdot)$. Since the generator $G_2(\cdot)$ is a pre-trained network, its parameter has already been trained to approach a suitable and relatively fixed value. Specifically, the training of generator networks focuses on utilizing BP and GD algorithms to update and optimize the learnable parameter Θ_{G_1} of $G_1(\cdot)$. The joint loss function of the generator networks is given by

$$\min \mathcal{L}_G = \mathbb{E}[\ln(1 - D(G_2(G_1(\mathbf{y}))))] + \lambda \|\theta - G_2(G_1(\mathbf{y}))\|_2^2. \quad (22)$$

Over the process of repeated training, the judgment probability $D(\hat{\theta})$ obtained by inputting the generated localization vector $\hat{\theta}$ learned by the generator networks into the trained discriminator network gradually approaches one to the most extent, while reducing the MSE between θ and $\hat{\theta}$.

In the inference stage, the aggregated received signal vector \mathbf{y} is input into the trained generator networks for simple feed-forward calculation. Then, the generated localization vector representing the localization results can be obtained, and thus the multi-target simultaneous high-precision positioning can be achieved.

VI. PERFORMANCE EVALUATION

In this section, the theoretical performance bound of the positioning accuracy is analyzed. Specifically, the CRLB of the positioning task in the indoor VLP scenario is derived. The CRLB is an unbiased estimator widely adopted to evaluate the theoretical lower bound of positioning [40].

Theorem 1. Suppose that the noise ω_n corresponding to the visible light channels between each LED and each robot follows an i.i.d. Gaussian distribution of $\mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I}_{MK})$. The CRLB of the estimated distance $\hat{\mathbf{d}}$ in the positioning task is given by

$$\begin{aligned} \mathbb{E}[\|\hat{\mathbf{d}} - \mathbf{d}\|_2^2] &\geq \frac{1}{N_p} \\ &\left(\frac{2\pi\sigma_n}{(m+1)(m+3)A_{\text{eff}}G_{\text{filter}}G_{\text{conc}}\delta^{m+1}} \right)^2 \sum_{i=1}^M \sum_{j=1}^K (d_{ij})^{2m+8}. \end{aligned} \quad (23)$$

Algorithm 1: The Proposed GAN Enabled Multi-Target VLP (GAN-MVLP) Scheme: Training Stage

Input: Minibatches of training data set $\{(\mathbf{y}^q, \mathbf{H}^q, \theta^q)\}_{q=1}^Q$ of size- Q .
Input: Learning rate γ_G for generator network; Learning rate γ_D for discriminator network.
Input: Maximum number of iteration steps W and stepsize λ .

- 1 Initialize the network parameters Θ_{G_1} and Θ_D
- 2 Repeat
- 3 **for** $q = 1, 2, 3, \dots, Q$ **do**
- 4 Generate the latent channel matrix $\hat{\mathbf{H}} = G_1(\mathbf{y})$
- 5 Generate the generated localization vector $\hat{\theta} = G_2(G_1(\mathbf{y}))$
- 6 **for** $w = 1, 2, 3, \dots, D-1$ **do**
- 7 Calculate the mean square error $\|\theta - G_2(G_1(\mathbf{y}))\|_2^2$ and optimize the generated localization vector in a gradient descent manner

$$\hat{\theta}_{w+1} = \hat{\theta}_w - \lambda \frac{\partial}{\partial \hat{\theta}_w} (\|\theta - G_2(G_1(\mathbf{y}))\|_2^2)$$
- 8 **end**
- 9 **end**
- 10 Evaluate the loss \mathcal{L}_G and \mathcal{L}_D of the generator and discriminator given by $G_1(\cdot)$ and $D(\cdot)$
- 11 Update the parameters of the both networks

$$\Theta_{G_1} = \Theta_{G_1} - \gamma_G \frac{\partial}{\partial \Theta_{G_1}} \mathcal{L}_G; \Theta_D = \Theta_D - \gamma_D \frac{\partial}{\partial \Theta_D} \mathcal{L}_D$$
- 12 Until reaching the maximum training steps

Output: Trained parameters Θ_{G_1} and Θ_D

Proof. According to formula (3), the channel gain between M LEDs and K intelligent robot targets with PDs is given by

$$h_{ij} = \frac{(m+1) A_{\text{eff}} \cos^m(\alpha) \cos(\varphi) G_{\text{filter}} G_{\text{conc}}}{2\pi d^2}. \quad (24)$$

Assume that the vertical distance between the LED and the PD at the robot target is δ , then equation (24) can be rewritten as

$$h_{ij} = \frac{(m+1) A_{\text{eff}} \delta^{m+1} G_{\text{filter}} G_{\text{conc}}}{2\pi d^{m+3}}. \quad (25)$$

Note that the representation above is only a theoretical result without taking noise into consideration. In the actual operation process, considering the impact of the background noise on the visible light transmission links, the channel impulse response matrix is given by

$$\mathbf{h}_N = \mathbf{h} + \boldsymbol{\omega}_\Lambda, \quad (26)$$

where $\boldsymbol{\omega}_\Lambda$ is the estimation error of the CIR, which can be modeled by the additive white Gaussian noise with a distribution of $\mathcal{N}(\mathbf{0}, \sigma_\Lambda^2 \mathbf{I}_{MK})$. The probability distribution function of \mathbf{h}_N conditioned by distance $\mathbf{d} = [d_{11}, d_{21}, \dots, d_{MK}]$ can be expressed as

$$p_{\mathbf{h}_N | \mathbf{d}}(\mathbf{h}_N; \mathbf{d}) = \frac{1}{(2\pi\sigma_\Lambda^2)^{M/2}} \exp \left\{ -\frac{1}{2\sigma_\Lambda^2} \|\mathbf{h}_N - \mathbf{h}\|_2^2 \right\}. \quad (27)$$

Then, the Fisher information matrix of the distance \mathbf{d} can be obtained through the conditional probability distribution

function in (27), which is given by

$$\begin{aligned} \mathbf{Q}_d &\triangleq \mathbb{E} \left[\frac{\partial \ln(p_{h_N|d}(h_N; d))}{\partial d} \left(\frac{\partial \ln(p_{h_N|d}(h_N; d))}{\partial d} \right)^T \right] \\ &= \mathbb{E} \left[\frac{\partial h}{\partial d} \frac{\partial \ln(p_{h_N|d}(h_N; d))}{\partial h} \left(\frac{\partial h}{\partial d} \frac{\partial \ln(p_{h_N|d}(h_N; d))}{\partial h} \right)^T \right] \\ &= \mathbf{F} \mathbf{Q} \mathbf{F}^T, \end{aligned} \quad (28)$$

where \mathbf{F} is a matrix of size $MK \times MK$, which is given in detail in equation (29). To be more specific, we show the details of the first row of \mathbf{F} in (29) as an example which is given by

$$\left(\frac{\partial h_{11}}{\partial d_{11}} \dots \frac{\partial h_{M1}}{\partial d_{11}} \frac{\partial h_{12}}{\partial d_{12}} \dots \frac{\partial h_{M2}}{\partial d_{12}} \dots \frac{\partial h_{MK}}{\partial d_{1K}} \right). \quad (30)$$

The matrix \mathbf{Q} in (28) is given by

$$\begin{aligned} \mathbf{Q} &= \mathbb{E} \left[\frac{\partial \ln(p_{h_N|d}(h_N; d))}{\partial h} \left(\frac{\partial \ln(p_{h_N|d}(h_N; d))}{\partial h} \right)^T \right] \\ &= (\sigma_\Lambda^2 \mathbf{I}_{MK})^{-1}. \end{aligned} \quad (31)$$

Substituting equations (29) and (31) into (28), \mathbf{Q}_d can be deduced and it is a diagonal matrix, whose ij -th diagonal entry is given by

$$[\mathbf{Q}_d]_{ij,ij} = \left(\frac{(m+1)(m+3)A_{\text{eff}}G_{\text{filter}}G_{\text{conc}}\delta^{m+1}}{2\pi\sigma_\Lambda(d_{ij})^{m+4}} \right)^2. \quad (32)$$

Therefore, the CRLB of the unbiased estimator \hat{d} for locating multiple intelligent robot targets can be derived from the inverse of the Fisher information matrix, which is given in detail in (33).

According to the Corollary 1 in the work of T. Wei *et al* [26], the approximate lower bound of the estimation error ω_Λ is given by

$$\sigma_\Lambda^2 \geq \frac{\sigma_n^2}{N_p}, \quad (34)$$

where σ_n^2 denotes the variance of the Gaussian distributed noise for the visible light channels between the LEDs and the robot targets; N_p is the number of pilot subcarriers. Finally, the CRLB of the positioning task can be deduced by substituting (34) into (33). \square

VII. SIMULATION RESULTS

In this section, the performance of the proposed schemes for simultaneous localization of multiple intelligent robots is evaluated through extensive simulations in an indoor factory environment as shown in Fig. 2. Specifically, extensive simulations have been conducted to verify the performance of the proposed CSM-VLP and CoCSM-VLP schemes, as well as the proposed GAN-MVLP method. For the simulation setup, a factory workshop with size of $4 \times 4 \times 3 \text{ m}^3$ is considered, where sixteen LEDs on the ceiling are evenly deployed according to the system model in Section III. The plane for the robots to be deployed is divided into 400 grid points. In this way, the

distance between adjacent grid points is 0.2m. Meanwhile, considering the volume of the intelligent robot, a relatively sufficient space is reserved with this configured step size of the grid points. As for the parameters of the VLC transmitters and receivers, the half-power angle of LEDs is 60° and the effective detection area of PDs is 1cm^2 .

A. Performance of Positioning Accuracy

Firstly, in order to visually compare the performance of the three proposed schemes in locating multiple robot targets, Fig. 4 shows the performance of positioning accuracy for the proposed schemes of multi-target VLP. It can be observed from Fig. 4 that all the three proposed schemes can achieve relatively accurate positioning performance for the task of simultaneous localization of four targets. In detail, the CSM-VLP scheme only has a larger positioning error for a certain target compared to the other two schemes, while all the three proposed schemes can achieve a high-precision localization performance for the other robot targets. This indicates that when the number of robots to be located is small, the sparsity of the vector Θ is satisfied well, and all the three proposed schemes have good performance. It is also indicated from Fig. 4 that, the CoCSM-VLP scheme achieves a higher positioning accuracy than CSM-VLP thanks to the collaboration among distinct anchors through the cross correlation operation. With the help of this procedure, more geometric information and channel state information in the environment can be utilized for positioning.

In order to demonstrate the capability of the proposed GAN-MVLP scheme compared with classical CS-based CSM-VLP and CoCSM-VLP schemes, the number of robots to be located is further increased to eight for simulation. The result of the positioning performance is shown in Fig. 5, where the estimated positions obtained by different schemes and the corresponding actual positions are marked for comparison, and a short line connecting them is used to represent the positioning error. The estimated positions are labelled in the center of a grid point because the proposed schemes recover the grid points in which the targets are located, and thus its center coordinate is labelled on average. It can be seen that as the number of the intelligent robot targets increases to eight, the performance difference in positioning accuracy between the classical CS-based scheme and the GAN enabled scheme grows significant. Specifically, for some of the robot targets, the positioning performance of the classical CS-based CoCSM-VLP scheme is limited. The correct grid point may not be found successfully, but an adjacent grid point may be located instead. However, the GAN-MVLP scheme can accurately locate each robot target on the grid point that is closest to the target. This indicates that the proposed GAN-MVLP scheme can effectively learn the potential mapping relationships between the measurement data and the sparse locations, thereby further improving positioning accuracy, especially in the case of a large number of targets.

Considering that multiple targets are located simultaneously, the estimated positions of different targets may not necessarily have the same distance from their actual positions. For the convenience of statistics and representation, we use the average

$$\mathbf{F} = \begin{pmatrix} \frac{\partial h_{11}}{\partial d_{11}} & \dots & \frac{\partial h_{MK}}{\partial d_{1K}} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_{11}}{\partial d_{MK}} & \dots & \frac{\partial h_{MK}}{\partial d_{MK}} \end{pmatrix} = -\frac{(m+1)(m+3)A_{\text{eff}} G_{\text{filter}} G_{\text{conc}} \delta^{m+1}}{2\pi} \begin{pmatrix} (d_{11})^{m+4} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (d_{MK})^{m+4} \end{pmatrix}^{-1}. \quad (29)$$

$$\mathbb{E} [\|\hat{\mathbf{d}} - \mathbf{d}\|_2^2] \geq \text{tr} (\mathbf{Q}_{\mathbf{d}}^{-1}) = \left(\frac{2\pi\sigma_{\Lambda}}{(m+1)(m+3)A_{\text{eff}} G_{\text{filter}} G_{\text{conc}} \delta^{m+1}} \right)^2 \sum_{i=1}^M \sum_{j=1}^K (d_{ij})^{2m+8}. \quad (33)$$

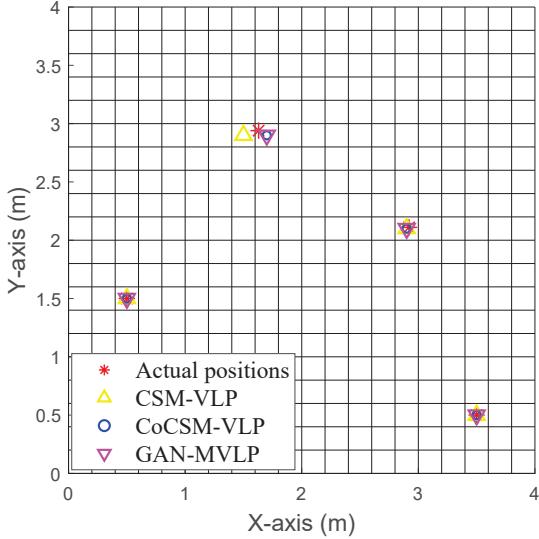


Fig. 4: Localization performance of three proposed schemes, i.e., CSM-VLP, CoCSM-VLP, and GAN-MVLP, for simultaneously locating four intelligent robot targets in a typical indoor factory scenario.

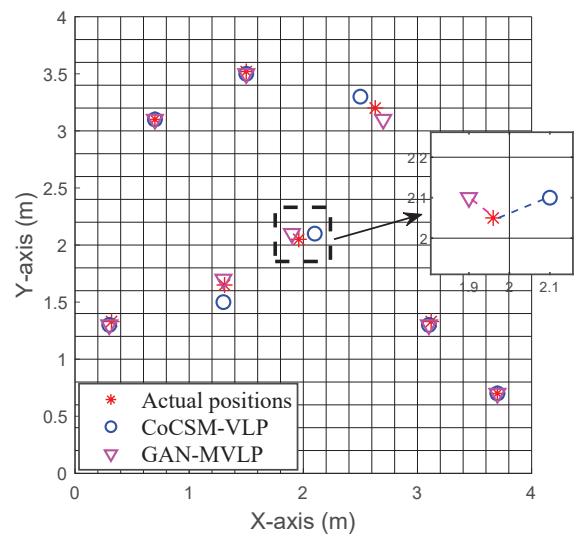


Fig. 5: Localization performance of the proposed GAN-MVLP scheme and the CoCSM-VLP scheme for simultaneously locating eight intelligent robot targets.

positioning error Δ to measure the accuracy of multi-target positioning, which is given by

$$\Delta = \frac{1}{K} \sum_{k=1}^K \left((x_k^{(\text{rx})} - x_k^{(\text{real})})^2 + (y_k^{(\text{rx})} - y_k^{(\text{real})})^2 \right)^{1/2}, \quad (35)$$

where $(x_k^{(\text{real})}, y_k^{(\text{real})})$, $k = 1, \dots, K$ denote the actual coordinates of the intelligent robots.

Next, in order to generalize the results obtained in Fig. 5 to a wider extent, we repeat substantial experiments under the same conditions. Based on the recorded data, a cumulative distribution function (CDF) with respect to the average positioning error is shown in Fig. 6. The CDF with respect to average positioning error can be utilized to evaluate the overall performance of different multi-target localization schemes. It can be observed that the three proposed schemes not only achieve high-precision simultaneous multi-target localization performance, but also significantly outperform the traditional RSS based method [42] in positioning accuracy. If we look at the target value of 0.9 for the cumulative distribution function, the average positioning errors of the proposed CSM-VLP, CoCSM-VLP and GAN-MVLP schemes are 15cm, 11cm and 6cm, respectively, which also verifies the advantage and

benefit introduced by cooperation and learning.

As indicated by Figs. 4 and 5, the positioning performance is closely related with the number of targets. In order to further investigate performance of the positioning accuracy of different schemes with respect to the number of robot targets, we have conducted simulations for the multi-target positioning task with different number of targets, and the results are reported in Fig. 7. Firstly, as an overall trend, it can be observed that the average positioning error of all schemes increases with the number of robot targets. The difference is that the three proposed schemes outperform traditional RSS-based VLP method [41] in terms of positioning accuracy, in the case of either a small or a large number of robots to be located. However, these three proposed schemes exhibit significant differences in their ability to address simultaneous localization of a large amount of targets. It is implied by the results in Fig. 7 that, both the CSM-VLP and CoCSM-VLP schemes based on traditional CS impose strong requirements on the sparsity of the spatial distribution of the targets with respect to the entity of the whole space, i.e., the sparsity of the localization vector Θ . Therefore, as the number of robots gradually increases, the positioning error will increase due to its impact on the spatial sparsity of the targets. However, with the help of inter-anchor cooperation via the cross-correlation

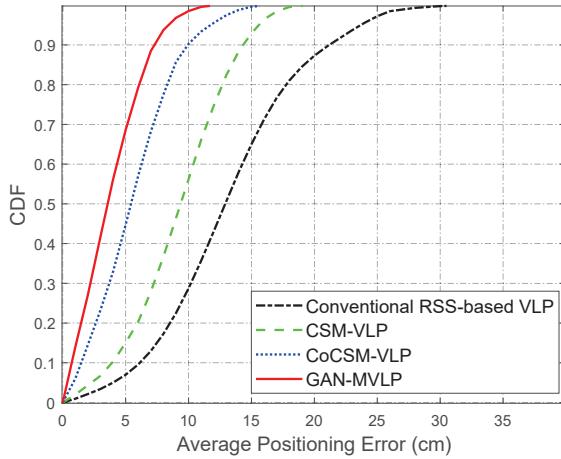


Fig. 6: CDF with respect to the average positioning error for the conventional RSS-based scheme and the three proposed schemes.

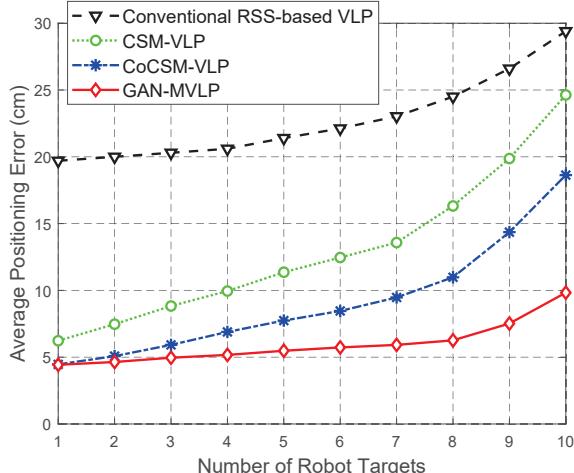


Fig. 7: The average positioning error of the three proposed schemes and the conventional RSS-based scheme with respect to the number of robot targets to be located.

operation, more available data can be utilized in the cross-correlation measurement model, so that the enhanced CoCSM-VLP scheme is less sensitive to the increase of the number of robots compared to the CSM-VLP scheme. Moreover, the GAN-MVLP scheme fully utilizes the strong ability of GAN in learning latent data distribution, so that it is observed that the average positioning error of GAN-MVLP does not change significantly with the increase of the number of targets. In particular, in the case of a large number of robots, the GAN-MVLP scheme can still maintain a relatively higher positioning accuracy.

B. Evaluation of Positioning Robustness

Next, in order to investigate the robustness of the proposed schemes in case of harsh environment, different SNR conditions are evaluated in simulation. Still, the average positioning error is used as a standard measure of the positioning perfor-

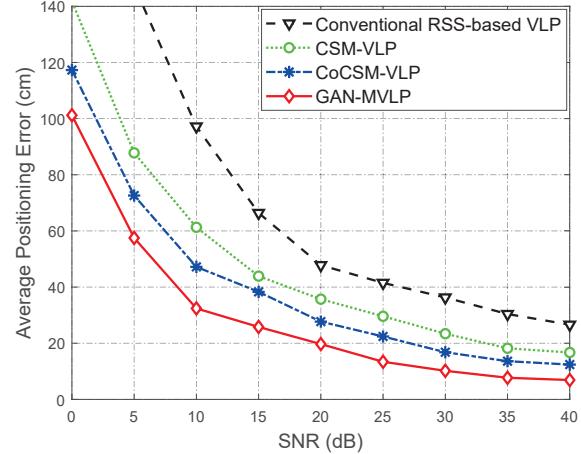


Fig. 8: Average positioning error with respect to the SNR for different VLP schemes for multi-target localization.

mance. Considering the scenario where there are eight robots to be located, the simulation results are shown in Fig. 8.

From Fig. 8, it can be observed that the proposed GAN-MVLP scheme can achieve centimeter-level multi-target positioning accuracy at the SNR of greater than 30dB, which is the best performance amongst these schemes evaluated. Meanwhile, the proposed CoCSM-VLP and CSM-VLP schemes can also reach an average positioning error smaller than 10cm at the SNR of 35dB and 40dB, respectively. When the SNR is small, a significant difference can be observed for different schemes in the robustness against strong noise. Among them, the GAN-MVLP scheme achieves the smallest average positioning error in the presence of intensive noise. It is also noted that the CoCSM-VLP scheme performs much better than the CSM-VLP scheme. Besides, it is clearly shown that all the three proposed schemes demonstrate better robustness to intensive noise than the conventional RSS-based VLP method.

VIII. CONCLUSION

In this paper, we introduced a compressed sensing (CS) based multi-target visible light positioning (VLP) framework designed to achieve high-precision simultaneous localization of multiple intelligent robots within indoor factory environments. This framework capitalizes on the sparsity of positional features in space, where the targets are confined to a limited number of positions in the plane. This characteristic enables us to transform the multi-target positioning problem into a sparse recovery problem. Our contributions include the development of a CS-based multi-target VLP (CSM-VLP) scheme, employing the aggregation of received signals and autocorrelation techniques to construct a measurement model. This scheme yielded effective localization results. To further enhance precision, we introduced an enhanced cooperative multi-target VLP (CoCSM-VLP) scheme. This scheme leverages inter-anchor cooperation, creating a cross-correlation measurement model that utilizes more environmental geometric information and channel state information to improve

localization. For scenarios involving a larger number of robots and heightened noise interference resistance, we devised a GAN model-based (GAN-MVLP) scheme. We also derived the theoretical boundaries of positioning accuracy. Simulation results affirm the superiority of our three proposed schemes over conventional RSS-based VLP methods in terms of positioning accuracy and robustness. At the CDF value of 0.9, the average positioning errors of the CSM-VLP, CoCSMVP and GAN-MVLP schemes are 15cm, 11cm and 6cm, respectively, which also verifies the advantage and benefit introduced by cooperation and learning. The proposed GAN-MVLP scheme can achieve centimeter-level multi-target positioning accuracy at the SNR of 30dB, which is the best performance amongst these schemes evaluated. This research may open promising avenues for advancing multi-target localization in indoor factory settings incorporated with intelligent robotics. In addition, the proposed multi-target VLP framework is easy to generalize and deployment in other location based service scenarios.

REFERENCES

- [1] K.-K. R. Choo, S. Gritzalis, and J. H. Park, "Cryptographic solutions for industrial internet-of-things: Research challenges and opportunities," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 8, pp. 3567–3569, Aug. 2018.
- [2] H. Kim, J. Kim, Y. Shin, J. Ahn, D. Kim, J. Park, B. Park, S. Huh, and S. Ahn, "Mitigation of frequency splitting phenomena using a matching capacitor in wireless power transfer system for automated guided vehicle," in *2020 IEEE Wireless Power Transfer Conference (WPTC)*, pp. 170–173, Seoul, Korea (South), Nov. 2020.
- [3] P. Das and L. Ribas-Xirgo, "A study of time-varying cost parameter estimation methods in automated transportation systems based on mobile robots," in *IEEE 21st International Conference on Emerging Technologies and Factory Automation (ETFA)*, Berlin, Germany, Sep. 2016.
- [4] Y. E. Song, M. Sik Kim, C. I. Kim, M. Hyung Song, G. S. Lee, J. Hwan Lim, and K. Lee, "Development of intelligent riding comfort monitoring system for automated vehicle: Conceptual structure and experiment design," in *2018 International Conference on Information and Communication Technology Robotics (ICT-ROBOT)*, Busan, Korea (South), Sep. 2018.
- [5] G. Schroeer, "A real-time UWB multi-channel indoor positioning system for industrial scenarios," in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Nantes, France, Sep. 2018.
- [6] C. Song, S. Liu, G. Han, P. Zeng, H. Yu, and Q. Zheng, "Edge-intelligence-based condition monitoring of beam pumping units under heavy noise in industrial internet of things for industry 4.0," *IEEE Internet of Things Journal*, vol. 10, no. 4, pp. 3037–3046, Feb. 2023.
- [7] C. Feng, W. S. A. Au, S. Valaee, and Z. Tan, "Received-signal-strength-based indoor positioning using compressive sensing," *IEEE Transactions on Mobile Computing*, vol. 11, no. 12, pp. 1983–1993, Dec. 2012.
- [8] A. Khalajmehrabadi, N. Gatsis, and D. Akopian, "Modern WLAN fingerprinting indoor positioning methods and deployment challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1974–2002, Thirdquarter 2017.
- [9] W. Zhao, S. Han, W. Meng, D. Sun, and R. Q. Hu, "BSDP: Big sensor data preprocessing in multi-source fusion positioning system using compressive sensing," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8866–8880, Sep. 2019.
- [10] L. E. M. Matheus, A. B. Vieira, L. F. M. Vieira, M. A. M. Vieira, and O. Gnawali, "Visible light communication: Concepts, applications and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3204–3237, Fourthquarter 2019.
- [11] Z. Ghassemlooy, S. Arnon, M. Uysal, Z. Xu, and J. Cheng, "Emerging optical wireless communications—advances and challenges," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 9, pp. 1738–1749, Sep. 2015.
- [12] T. Q. Wang, Y. A. Sekercioglu, A. Neild, and J. Armstrong, "Position accuracy of Time-of-Arrival based ranging using visible light with application in indoor localization systems," *Journal of Lightwave Technology*, vol. 31, no. 20, pp. 3302–3308, Oct. 2013.
- [13] B. Zhu, J. Cheng, Y. Wang, J. Yan, and J. Wang, "Three-dimensional VLC positioning based on angle difference of arrival with arbitrary tilting angle of receiver," *IEEE Journal on Selected Areas in Communications*, vol. 36, no. 1, pp. 8–22, Jan. 2018.
- [14] J. Wang, S. Kwon, P. Li, and B. Shim, "Recovery of sparse signals via generalized orthogonal matching pursuit: A new analysis," *IEEE Transactions on Signal Processing*, vol. 64, no. 4, pp. 1076–1089, Feb. 2016.
- [15] F. Y. Cai, Lei, Y. Xiang, T. Zhu, X. Li, and H. Zeng, "Fast compressed sensing recovery using generative models and sparse deviations modeling," in *IEEE International Conference on Visual Communications and Image Processing (VCIP)*, pp. 447–450, Macau, China, Dec. 2020.
- [16] E. Balevi, A. Doshi, A. Jalal, A. Dimakis, and J. G. Andrews, "High dimensional channel estimation using deep generative networks," *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 1, pp. 18–30, Jan. 2021.
- [17] M. Mardani, E. Gong, J. Y. Cheng, S. S. Vasanawala, G. Zaharchuk, L. Xing, and J. M. Pauly, "Deep generative adversarial neural networks for compressive sensing MRI," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 167–179, Jan. 2019.
- [18] X. Wang and S. Liu, "Cooperative visible light positioning: A compressed sensing based framework," in *IEEE International Conference on Communications (ICC)*, Rome, Italy, May. 2023.
- [19] Y. Laili, Z. Chen, L. Ren, X. Wang, and M. J. Deen, "Custom grasping: A region-based robotic grasping detection method in industrial cyber-physical systems," *IEEE Transactions on Automation Science and Engineering*, vol. 20, no. 1, pp. 88–100, Jan. 2023.
- [20] J. Xin, G. Xie, B. Yan, M. Shan, P. Li, and K. Gao, "Multimobile robot cooperative localization using ultrawideband sensor and GPU acceleration," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 4, pp. 2699–2710, Oct. 2022.
- [21] I. Okunovich and D. Trinitatova, "Deltacharger: Charging robot with inverted delta mechanism and CNN-driven high fidelity tactile perception for precise 3D positioning," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7604–7610, Oct. 2021.
- [22] Y. Yu, R. Chen, L. Chen, W. Li, Y. Wu, and H. Zhou, "H-WPS: Hybrid wireless positioning system using an enhanced Wi-Fi FTM/RSSI/MEMS sensors integration approach," *IEEE Internet of Things Journal*, vol. 9, no. 14, pp. 11827–11842, July 2022.
- [23] E. Gnilta, E. Lei, J. Langerman, H. Huang, and C. Studer, "CSI-based multi-antenna and multi-point indoor positioning using probability fusion," *IEEE Transactions on Wireless Communications*, vol. 21, no. 4, pp. 2162–2176, April 2022.
- [24] Y. Li, X. Hu, Y. Zhuang, Z. Gao, P. Zhang, and N. El-Sheimy, "Deep reinforcement learning (DRL): Another perspective for unsupervised wireless localization," *IEEE Internet of Things Journal*, vol. 7, no. 7, pp. 6279–6287, July 2020.
- [25] M. F. Keskin, S. Gezici, and O. Arikan, "Direct and two-step positioning in visible light systems," *IEEE Transactions on Communications*, vol. 66, no. 1, pp. 239–254, Jan. 2018.
- [26] T. Wei, S. Liu, and X. Du, "Visible light integrated positioning and communication: A multi-task federated learning framework," *IEEE Transactions on Mobile Computing (Early Access)*, pp. 1–18, 2022.
- [27] B. Zhou, A. Liu, and V. Lau, "Performance limits of visible light-based user position and orientation estimation using received signal strength under NLOS propagation," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5227–5241, Nov. 2019.
- [28] N. Garcia, H. Wymeersch, E. G. Larsson, A. M. Haimovich, and M. Coulon, "Direct localization for massive MIMO," *IEEE Transactions on Signal Processing*, vol. 65, no. 10, pp. 2475–2487, May 2017.
- [29] H. Jamali-Rad, H. Ramezani, and G. Leus, "Sparsity-aware multi-source RSS localization," *Signal Processing*, vol. 101, pp. 174–191, 2014.
- [30] X. Gong, J. Liu, S. Yang, G. Huang, and Y. Bai, "A usability-enhanced smartphone indoor positioning solution using compressive sensing," *IEEE Sensors Journal*, vol. 22, no. 3, pp. 2823–2834, Feb. 2022.
- [31] R. Zhang, W.-D. Zhong, K. Qian, S. Zhang, and P. Du, "A reversed visible light multtarget localization system via sparse matrix reconstruction," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 4223–4230, Oct. 2018.
- [32] T. Wei, S. Liu, and X. Du, "Learning-based efficient sparse sensing and recovery for privacy-aware IoT," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9948–9959, June 2022.
- [33] S. Liu, L. Xiao, L. Huang, and X. Wang, "Impulsive noise recovery and elimination: A sparse machine learning based approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 3, pp. 2306–2315, March 2019.

- [34] J. Gao, C. Zhong, G. Y. Li, and Z. Zhang, "Deep learning-based channel estimation for massive MIMO with hybrid transceivers," *IEEE Transactions on Wireless Communications*, vol. 21, no. 7, pp. 5162–5174, July 2022.
- [35] L. E. M. Matheus, A. B. Vieira, L. F. M. Vieira, M. A. M. Vieira, and O. Gnawali, "Visible light communication: Concepts, applications and challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3204–3237, Fourthquarter 2019.
- [36] D. Karunaratne, F. Zafar, V. Kalavally, and R. Parthiban, "LED based indoor visible light communications: State of the art," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1649–1678, Thirdquarter 2015.
- [37] Y. Wang, M. Chen, Z. Yang, T. Luo, and W. Saad, "Deep learning for optimal deployment of UAVs with visible light communications," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7049–7063, Nov. 2020.
- [38] J. W. Choi, B. Shim, Y. Ding, B. Rao, and D. I. Kim, "Compressed sensing for wireless communications: Useful tips and tricks," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1527–1550, Thirdquarter 2017.
- [39] T. Wei, S. Liu, and D. Su, "Generative adversarial network enabled sparse signal compression and recovery for internet of medical things," in *ACM UbiComp'21*, pp. 678–683, Electr Network, Sep. 2021.
- [40] S. Liu and D. Su, "Deep learning based underwater acoustic channel estimation exploiting physical knowledge on channel sparsity," in *ACM UbiComp'21*, pp. 655–659, Electr Network, Sep. 2021.
- [41] F. Yin, C. Fritzsche, F. Gustafsson, and A. M. Zoubir, "TOA-based robust wireless geolocation and cram-rao lower bound analysis in harsh LOS/NLOS environments," *IEEE Transactions on Signal Processing*, vol. 61, no. 9, pp. 2243–2255, May 2013.
- [42] N. Huang, C. Gong, J. Luo, and Z. Xu, "Design and demonstration of robust visible light positioning based on received signal strength," *Journal of Lightwave Technology*, vol. 38, no. 20, pp. 5695–5707, Oct. 2020.



Sicong Liu (Senior Member, IEEE) received the B.S.E. and Ph.D. degrees (Highest Hons.) in electronic engineering from Tsinghua University, Beijing, China, in 2012 and 2017, respectively. He is an Associate Professor with the Department of Information and Communication Engineering, School of Informatics, Xiamen University, Xiamen, China. He was a Senior Engineer with Huawei Technologies Company Ltd., China, from 2017 to 2018. He was a Visiting Scholar with the City University of Hong Kong in 2010. His current research interests are compressed sensing, AI-assisted communications, integrated sensing and communications, and visible light communications. He has authored over 60 journal or conference papers, and four monographs in the related areas.

Dr. Liu won the Best Paper Award at ACM UbiComp 2021 CPD WS as the corresponding author, and the Second Prize in the Natural Science Award of Chinese Institute of Electronics. He has served as the associate editor or TPC chair of several IEEE and other international academic journals and conferences. He is a Senior Member of IEEE and China Institute of Communications.



Jian Song (Fellow, IEEE) received the B. Eng. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1990 and 1995, respectively. He is currently the Director with Tsinghua DTV Technology R&D Center. Dr. Song has authored or coauthored more than 300 peer-reviewed journal and conference papers. He holds two U.S. and more than 80 Chinese patents. His research interest include the area of digital TV broadcasting, fiber-optic, satellite and wireless communications, and the power-line communications.

He is also the Fellow of IET.



Zhu Han (Fellow, IEEE) received the B.S. degree in electronic engineering from Tsinghua University, Beijing, China, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland, College Park, MD, USA, in 1999 and 2003, respectively.

He was a Research and Development Engineer with JDSU, Germantown, MD, USA, from 2000 to 2002. He was a Research Associate with the University of Maryland from 2003 to 2006. He was an Assistant Professor with Boise State University, Boise, ID, USA, from 2006 to 2008. He is currently a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department as well as with the Computer Science Department, University of Houston, Houston, TX, USA. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid.

Dr. Han received the NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the Journal on Advances in Signal Processing in 2015, the IEEE Leonard G. Abraham Prize in the field of Communications Systems (Best Paper Award in IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS) in 2016, and several best paper awards in IEEE conferences. He is also the Winner of the 2021 IEEE Kiyo Tomiyasu Award, for outstanding early to mid-career contributions to technologies holding the promise of innovative applications, with the following citation: for contributions to game theory and distributed management of autonomous communication networks. He has been a 1% Highly Cited Researcher since 2017 according to the Web of Science. He was an IEEE Communications Society Distinguished Lecturer from 2015 to 2018, and has been an AAAS Fellow since 2019 and an ACM Distinguished Member since 2019.



Xianyao Wang received the B.S. degree in communication engineering from Ningxia University, Yinchuan, China in 2021. He is currently pursuing the M.S. degree with the Department of Information and Communication Engineering, Xiamen University, Xiamen, China. His research interests include compressed sensing and visible light positioning.