BEST-RIM: a mmWave Beam Steering Approach based on Computer Vision-enhanced Reconfigurable Intelligent Metasurfaces

Valeria Loscri, Carola Rizza, Abderrahim Benslimane, Anna Maria Vegni, Eros Innocenti, Romeo Giuliano

BEST-RIM: a mmWave **Be**am **St**eering Approach based on Computer Vision-enhanced **R**econfigurable **I**ntelligent **M**etasurfaces

Valeria Loscrí, *Senior Member, IEEE,* Carola Rizza, *Member, IEEE,* Abderrahim Benslimane, *Senior Member, IEEE,* Anna Maria Vegni, *Senior Member, IEEE,* Eros Innocenti, *Member, IEEE,* Romeo Giuliano, *Senior Member, IEEE*

***Abstract*—In this work, we consider Reconfigurable Intelligent Meta-surfaces (RIMs) to provide effective beam steering functionality, while enhancing the coverage of the fifth generation (5G) users in a urban city mobility context. In particular, we enable communications in the millimeter wave (mmWave) frequencies, that can occur simultaneously with other 5G communications. The main objective of this paper is to demonstrate the feasibility and advantages of an integrated Computed Vision (CV) approach for tuning the meta-atom states of a RIM structure working at mmWave frequency band, without estimating the interferes' radio channel. In particular, the CV system will feed a logic unit running a Machine Learning (ML) algorithm to compute in real-time the coding schemes, namely the sequence of binary states associated to each unit-cell of the RIM, in order to obtain the target reconfigurable radiation pattern. Specifically, a Genetic Algorithm (GA) is introduced to derive the most suitable radiation pattern for Beam Steering application, based on the input of a CV infrastructure. Results show the feasibility of such a kind of system, with an higher coverage achieved in dense scenarios, by improving the robustness against the potential blockages introduced by the mmWave technology. Moreover, we demonstrate the system is robust against the inaccuracy introduced by the CV and GA.**

***Index Terms*—Reconfigurable Intelligent Meta-surfaces, Machine Learning, Computer Vision, mmWave, Beam Steering.**

I. Introduction

In recent years, mmWave technology has acquired an increasing interest to be exploited in future wireless cellular communications, such as in the sixth generation (6G) network scenarios [1], [2], [3]. Due to the large bandwidth, very high data rates in the order of gigabit per second (Gbps) are expected to be achieved with much increased capacity, as compared to other existing wireless communication systems. Furthermore, since mmWave links present high directionality feature, an efficient spatial and frequency reuse can be obtained, with the possibility to exploit Line-of-Sight (LOS) links that are very beneficial for interference control [4].

However, some important issues related to the specific propagation features of mmWave signal need to be considered and addressed in order to exploit the enormous potentiality of this technology. Among the main challenges with mmWave we remind the significant path loss that the mmWave signals experience [5], [6], the blockage due to obstacles, the atmospheric attenuation, as well as harsh weather conditions such as rain, absorption by water vapors, and oxygen, that degrade the mmWave channel.

Recently, in order to address the main issues of mmWaves technology, the concept of Reconfigurable Intelligent Meta-Surface has been introduced in the research community [7], [8]. RIMs are based on passive materials, aimed to overcome the LOS blockage, to improve the communication link by effectively reconfiguring the propagation of the electromagnetic signal. The wireless signal can be accordingly manipulated in order to define a given reflecting direction, by adjusting the phase shift and/or the amplitude of each element comprising the RIM *i.e.*, meta-atom [9]. RIM technology can then allow the definition of the Smart Radio Environment (SRE) concept [10], [11] *i.e.*, the surrounding environment is enriched with intelligent devices that control the wireless propagation towards desired directions. This approach can be applied both in indoor and outdoor scenarios, and is useful not only to focus the wireless signal towards a final destination, but also to put the propagation nulls towards unwanted/undesired directions, mainly for safety applications.

The use of RIMs in the mmWave band has been adopted in [12], where a RIM-assisted approach for efficient and green resource allocation is presented. In [13], the authors have proposed a solution for indoor mmWave communication, even in presence of blockage, validating the results with a test-bed. However, most works on the use of RIMs for wireless communications mainly focus on the conventional signal processing aspects. It has been demonstrated that the joint use of Artificial Intelligence (AI) algorithms [14] with RIM-based systems can provide enhanced performance, and fully fits within the context of SRE [15].

Concerning the optimization of RIM-based systems, previous contributions in literature have regarded the feasibility to achieve beamforming based on RIM exploitation, by proposing optimization approaches such as in [16], where the authors formulate a joint optimization problem for optimizing active and passive beamforming. Their approach is based on an ideal case study, with continuous phase shifts. The discrete phase shift is considered in [17]. Another algorithm aiming to

Valeria Loscrí and Carola Rizza are with the FUN Team at Inria Lille - Nord Europe

Abderrahim Benslimane is with University of Avignon, France.

Anna Maria Vegni is with the Dept. of Industrial, Electronics and Mechanical Engineering at Roma Tre University.

Eros Innocenti and Romeo Giuliano are with the Dept. of Engineering Science Guglielmo Marconi University

Manuscript received XXX, XX, 2015; revised XXX, XX, 2015.

maximize the transmit beamforming is proposed in [18]. By exploiting a jointly optimization of the precoding strategies of the Base Station and the RIM parameters, the authors demonstrate the minimization of the system sum mean squared error (SMSE). All these contributions are based on active channel control, which incurs in a very high overhead limiting its efficiency and can hinder the real-time responsiveness of certain applications as the one considered in this work.

Other approaches have been proposed to relax the associated feedback overhead and are mostly based on statistical Channel State Information (CSI) [19], [20], [21], but still rely on a control channel to steer the RIM beams. The main rationale for our work is to relax the control channel and the high overhead associated to, by exploiting a pre-existing and well-deployed infrastructure, namely the cameras that are nowadays massively deployed in the modern cities. Other works such as [22] and [23] exploit the available cameras for object detection, thus supporting drivers and impaired people. Nevertheless, to the best of our knowledge, this is the first work addressing mmWave beamforming based on RIM, jointly considering the CV.

Among the AI-based approaches, a very interesting concept that has started to capture the interest of scientific community in a synergic way with the wireless networks and with mmWave technology, is the Computer Vision (CV) technology. CV permits to leverage visual data originated by different vision sensors (*i.e.*, cameras), and can be regarded as a key enabler technology for Beyond-5G wireless networks [24]. An early contribution on this topic proposes the combination of CV-based techniques with wireless networks, in particular with millimeter-wave radar [25]. The authors demonstrate how images can be leveraged for improving the distinction between lane lines at a short distance. In [26] authors propose the integration of CV with wireless sensors to perform a high accurate indoor localization, through the use of Wi-Fi signals emitted by the smartphones for fine users' positioning. An interesting perspective on the CV exploitation is also presented in [24], where the authors introduce the concepts of *View to Communicate (V2C)* and *Communicate to View (C2V)* paradigms. In V2C, the visual data are used to enhance the performance of an unknown wireless channel, by accounting for blockages that can occur in a data transmission link. In C2V, the opposite approach is considered.

Based on the recent encouraging results obtained from the integrated use of (*i*) CV algorithms, and (*ii*) RIM technology for a wireless communication system [27], in this paper we consider this joint combination to be applied in a wireless network context. In respect of [27], in this work we detail the modeling of the users and the RIMs, and we provide a detailed analysis of the performance of the systems by considering mmWave users. The main objective is to exploit the information provided by the CV to compute optimal radiation patterns for enhancing the communication between users in mmWave links. The optimality is based on the radiation pattern configuration, that based on the user's position maximize the received power. Specifically, we consider a V2C approach based on the joint use of CV and RIM devices. To the best of our knowledge, this is the first work presenting a synergistic approach of the CV technology with the RIM paradigm in a wireless context, aiming to improve the system performance, as well as to reduce interference to the network nodes. We propose a complete framework, namely BEST-RIM –a mmWave Beam Steering Approach based on Computer Vision-enhanced RIMs–, based on the integration of a CV sub-system with a RIM-based one. The proposed framework exploits the CV technology in order to enhance mmWave cellular communications through RIM nodes. Videos collected by cameras are processed and used to "understand" the traffic nature of a given scenario. The CV sub-system acts as an "observer" of the wireless scenario (*e.g.*, mobile users walking in an urban area), capturing and processing video frames in real-time. The output of the CV-subsystem, expressed as azimuth and zenith angles, as well as the distances between a RIM node and a generic user, triggers the RIM-based subsystem. All these data will feed a Processing Unit (PU) running a GA for computing the coding schemes in a dynamic way, namely the states associated with the single meta-atoms, in order to generate a desired radiation pattern at the RIM. The choice of the desired radiation pattern occurs with that one maintaining the maximum coverage of a mobile user *i.e.*, guaranteeing the connectivity link.

Previous literature on RIM has been mostly based on the optimization problem of the meta-surface, relying on the information of the channel state or considering statistical CSI. To the best of our knowledge, there are no global RIM-based architectures or a complete system, encompassing the generation of the input, their treatment and the estimation of the performance based on the whole chain. Differently, this paper aims to propose a RIM-based framework for effective beamsteering functionality, supported by a CV-subsystem feeding a GA with the information about interference in a given scenario. Indeed, thanks to the BEST-RIM features, it is not necessary to estimate the interferers' radio channel, since we obtain the information for the GA, directly from the CV system.

The main contributions of this paper can be summarized as follows:

- We exploit the concept of V2C by means of a CV-based approach to monitor traffic load in a real network scenario. The CV solution is introduced together with a RIM-based sub-system, performing beam steering functionality in the mmWave working frequency;
- We present the proposed BEST-RIM framework, and a detailed description of main entities is provided;
- We design and evaluate a GA approach, fed by the output from the CV sub-system, in order to compute the optimal coding schemes of a full reconfigurable structure, namely the sequence of states associated to the unit cells for maximizing the receiver power;
- We explicitly account the interference generated by mmWave users and we derive the coverage probability, expressed as the probability that a certain receiver node is reached by a RIM through the Complement Cumulative Distributed Function;
- We evaluate the impact of the error for both the GA and the CV sub-system, in terms of coverage probability.

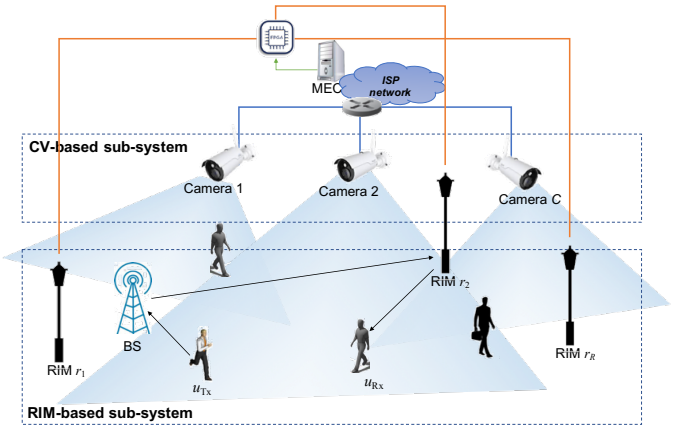

Fig. 1. BEST-RIM framework. The RIM device is trigged by CV data *i.e.*, locations and Direction of Arrival (DoA) angles, derived by the CV-based sub-system. The RIM device is able to derive the best configuration for communication with the receiver target node, and to reduce the interference affecting other data links.

The rest of the paper is organized as follows. In Section II, we present the overall system model. In Section III, we describe the GA, that is a specific ML algorithm implemented in the PU. Section IV is dedicated to the CV approach. In Section V, we provide the performance evaluation of each module of the overall system, and based on these results, we provide the evaluation of the overall system. Finally, conclusions are drawn at the end of this paper, followed by a final discussion.

II. System Model

In this paper, we focus on beam steering applications on mmWave communications in heterogeneous networks, and we aim to characterize the interference generated by other transmitters [28], [29]. We assume a hybrid system, with cellular communications and communications between RIM-coated objects, acting as both smart reflectors, and user devices. Notice that nodes can transmit and receive, based on the traditional cellular spectrum or by exploiting mmWave spectrum.

Let us consider an outdoor scenario as depicted in Fig. 1, comprised of a set $\mathcal{U} = \{u_1, u_2, \ldots, u_U\}$ with $U \in \mathbb{Z}^+$, of mobile nodes (*i.e.*, pedestrians) each of them equipped with a personal device. We also assume the set $\mathcal{R} = \{r_1, r_2, \ldots, r_R\}$ with $R \in \mathbb{Z}^+$, of RIM-coated devices, each of them deployed in a fixed position and fully integrated in the network scenario (*e.g.*, a RIM can coat the top of a street lamp or a statue in the middle of an urban place). Finally, the set $\mathcal{C} = \{c_1, c_2, \ldots, c_C\}$ with $C \in \mathbb{Z}^+$ collects the camera nodes, installed in main city routes. Each camera is assumed to be wired-connected to the Mobile Edge Computing (MEC) node for transmitting visual content to be further elaborated. All the details of the BEST-RIM network framework are depicted in Fig. 1.

As observed, BEST-RIM framework is comprised of (*i*) a CV-based sub-system and (*ii*) a RIM-based sub-system. The CV-based sub-system is comprised of camera devices installed in main city squares and routes. It aims to monitor a given area, where pedestrians are moving. Each camera can store visual content, which is directly processed by a video processor, able to detect people shapes and then computes the distances from a reference point. Specifically, we assume that the visual content captured by the cameras is transmitted to a MEC node for further processing. For deployment constraints and for scale costs, the MEC device is assumed to be located in the Internet Service Provider (ISP) network.

The logic flow behind the BEST-RIM framework is depicted in Fig. 2, where we assume a source node (*i.e.*, $u_{Tx} \in \mathcal{U}$) aiming to transmit to a receiver node (*i.e.*, $u_{Rx} \in \mathcal{U}$). The communication links are from node u_{Tx} to a mmWave Base Station (BS) (see arrow #1), and from the BS to a given RIM deployed in the reference area (see arrow #2). The RIM is able to redirect the received impinged signal toward the receiver node u_{Rx}, accordingly trigged by the FPGA node. The latter is able to control the RIM, which, thanks to the GA, is able to find the best radiation pattern (*i.e.*, final beam) for communication with the receiver node (see arrow #3) and reduce the interference on other neighboring transmissions. In Fig. 2, the CV elaboration is used to derive the information in terms of distances between the users and the RIM, and the angles between a RIM and the receiver node u_{Rx}. The CV-processed data are used as input for an intelligent logic *e.g.*, a unit integrating a learning approach. In this work, our learning approach is based on a GA approach, able to compute the best configuration of the single meta-atoms of the meta-surface, in order to generate its optimal configuration in terms of radiation pattern, to maximize the received power. For sake of simplicity, we assume that all the nodes use the same wireless technology for transmission and then the information about the monitoring area are derived based on the distance computed by applying a propagation model related to this specific technology. Furthermore, the u_{Rx} target receiver is known by the RIM device and the RIM radiation pattern configuration will be computed in order to maximize the power received by the receiver node, thus putting nulls to neighboring interferents.

More specifically, the User Equipment (UE) of the u_{Tx} forwards data to the eNB (*i.e.*, the BS in the figure). The eNB/BS transmits the information to the RIM, that will also receive information by the CV. The Unit Logic (UL), integrating the ML approach (*i.e.*, the GA in this case), will elaborate the data received both by the CV and the transmitter for the target receiver (*i.e.*, u_{Rx}) and the correct sequence code of the RIM, will be coded through the FPGA, in order to generate the corresponding target radiation pattern.

Following the scenario in Fig. 2, we assume the RIM devices are randomly distributed according to a Poisson Point Process (PPP) distribution Φ_{RIM} with a certain density λ_{RIM} [30]. Also the set of users $\mathcal{U}$ are randomly and independently distributed, following a PPP distribution characterized as:

$$\Phi_{\mathcal{U}} = \{(L, D, P, \psi, C)\}, \tag{1}$$

where $\{L\}$ is the set of the locations of the users distributed as an unmarked PPP with associated intensity λ, $\{D\}$ is the set of the distances between the j-th RIM node and the i-th user *i.e.*,

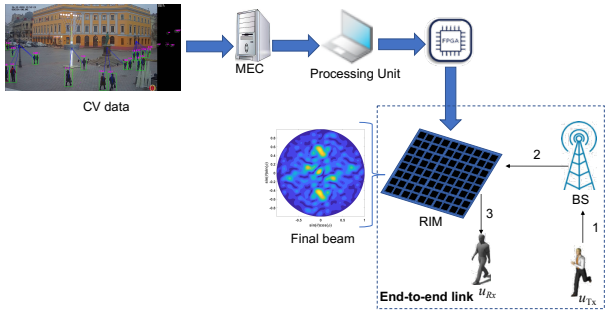

Fig. 2. Schematic of BEST-RIM approach. The RIM object is trigged by CV data (*i.e.*, estimated locations and angles) derived by the CV-based sub-system. The RIM device is able to derive the best configuration for communication with the receiver target node, and to reduce the interference affecting other data links.

$D = \{d_{11}, d_{12}, \ldots, d_{ji}, \ldots, d_{RU}\}$. The set of the distances $\{D\}$ follows a Rayleigh distribution with a Probability Density Function (PDF), defined as

$$f_D(x) = 2\pi d_{ji} x \exp(-d_{ji}\pi x^2), \tag{2}$$

with d_{ji} [m] representing the distance between the j-th RIM node and the i-th receiver node. Based on this assumption, it follows that a generic mmWave receiver is randomly located close to a RIM following a Gaussian distribution [31].

In Eq. (1), $\{P\}$ is the set of the transmitting powers from the j-th RIM to the i-th users *i.e.*, $P = \{P_{11}, P_{12}, \ldots, P_{ji}, \ldots, P_{RU}\}$, and $\{\psi\}$ is the set of the oriented angles between the j-th RIM node and the i-th user to the x-axis, assumed as equiprobably distributed. Finally, $\{C\}$ is the set of $i.i.d$ random variables following a Bernoulli distribution, and represents the type of communication adopted by the i-th user, namely traditional communication (*i.e.*, $C_i = 0$) and mmWave (*i.e.*, $C_i = 1$). Specifically, we assume that the i-th node selects the mmWave communication mode (*i.e.*, $C_i = 1$), when the distance $d_{ji} < \chi$, with χ [m] as a threshold of the distance between the j-th RIM and the i-th receiver.

Based on these assumptions, we consider the set $\{\Phi^{(mmW)}\}$, constituted by the pairs of the RIMs connected to the receiver nodes, communicating in mmWave, is PPP distributed with the following expression for the intensity:

$$\begin{aligned}\lambda^{(mmW)} &= \lambda \Pr(C_i = 1)\Pr(d_{ji} < \chi) = \\ &= \lambda \Pr(C_i = 1)(1 - e^{-d_{ji}\pi\chi^2}).\end{aligned} \tag{3}$$

A. The Interference Model

Let us consider the mmWave LOS link between the j-th RIM transmitting to the i-th receiver node. Following the usual modeling of channels considered in the literature [32], the received signal can be written by considering two components, the "desired" signal and the interference as follows:

$$\begin{aligned}Y_i = &\sqrt{P_{ji} d_{ij}^{-\alpha} g_{Tx,j} g_{Rx,i} h_{ji}} V_i + \\ &+ \sum_{L_\ell \in [\Phi_{\mathcal{U}} - L_i]} \sqrt{P_j \| L_\ell - L_i \|^{-\alpha} g_{Tx,\ell} g_{Rx,i} h_{\ell i}} V_j + \\ &+ Z,\end{aligned} \tag{4}$$

where L_i and L_ℓ are the positions of the i-th receiver and the ℓ-th interferent, with $\ell \neq i$, and $j \in \mathbb{Z}^+$, respectively; d_{ij} denotes the distance between nodes i and j, and P_{ji} is the transmitting power from the j-th RIM towards the i-th user. It is worth to outline that the j-th RIM is constituted by quasi-passive components, and it does not amplify the impinging signal originated by the BS, since it is used as a smart reflector. The threshold value χ is also established on the basis of the distance between the j-th RIM and the i-th receiver node, by considering these specific characteristics of the RIMs. In Eq. (4), P_j denotes the transmitting power of the j-th node, the parameter α is the path loss coefficient, $g_{Tx,j}$ and $g_{Rx,i}$ are the antenna gains of the j-th RIM and the i-th receiver node, respectively; h_{ji} represents the channel fading between the i-th receiver and the j-th RIM reflecting the signal, while $h_{\ell i}$ is the channel fading of the ℓ-th interfering node on the i-th receiver node. Finally, V_i is the unit-variance signal at the i-th receiver node, and Z denotes the additive white Gaussian noise. Notice that, according to recent results in [33], we consider the channel propagation modeled according to a Rayleigh fading, and so we can approximate $h_{ji} \sim \exp(1)$, and $h_{\ell i} \sim \exp(1)$, and also fading is independent over space.

The radiation pattern is derived by the total reflected electric field, which is computed through the Huygens formula, shown in Eq. (5). The x-, y- and z- components of the electric field are derived from its spherical expression $E(\theta, \phi)$ and will be used for estimating the transmitting power P_{ji} in Eq. (4) *i.e.*,

$$\begin{aligned}E(\theta,\phi) = \sum_{m=1}^{M}\sum_{n=1}^{N} &A_{mn} e^{j\alpha_{mn}} f_{mn}(\theta_{mn}, \phi_{mn}) \ \Gamma_{mn} e^{j\Phi_{mn}} \\ &f_{mn}(\theta,\phi) e^{jk_0 \zeta_{mn}(\theta,\phi)},\end{aligned} \tag{5}$$

where M and N are the number of meta-atoms in the two dimensions, A_{mn} and α_{mn} are the amplitude and phase of the incident wave on the (mn)-th meta-atom, θ_{mn} and ϕ_{mn} are the elevation and azimuth angles of the source relative to the (mn)-th meta-atom, respectively. Furthermore, Γ_{mn} and Φ_{mn} are the amplitude and the phase reflection coefficients, respectively, and f_{mn} represents the scattering pattern of the (mn)-th meta-atom that comprises the RIM. We assume that $f_{mn}(\theta, \phi) = \cos\theta$, [34]. Finally, $\zeta_{mn}(\theta, \phi)$ is a relative phase shift and k_0 equal to $2\pi/\lambda$, with λ [m] the wavelength of the transmitted signal.

B. The SINR Characterization

In order to characterize the performance of the hybrid communication system, based on both "traditional" (*e.g.*, at frequencies lower than 6 GHz) and mmWave communication, we need to derive the interference at the i-th receiver node. Based on the previous assumption, the interference for a node is only based on the mmWave communication links from other nodes.

Let us assume $L_i = (x_i, y_i)$ as the location of the i-th receiver, and the total interference $\mathcal{I}$ can be derived as:

$$\mathcal{I} = \sum_{L_j \in \Phi^{(mmW)}, j \neq i} h_{ji} P_j g_{Tx,j}(\theta(j,i), \phi(j,i)) g_{Rx,i}(\theta(i,j), \phi(i,j)) L_{ij}^{-\alpha}, \tag{6}$$

where we consider the interference of a user by accounting for its specific position through the gains $g_{Tx,j}(\theta(j,i),\phi(j,i))$ and $g_{Rx,i}(\theta(i,j),\phi(i,j))$. By applying the Slivnyak's Theorem [35], [36], the Signal-to-Interference plus Noise Ratio (SINR) can be derived as:

$$SINR^{(mmW)} = \frac{P_i}{\mathcal{I} + P_{N,i}}, \tag{7}$$

where P_i [W] is the power at the i-th receiver, $P_{N,i}$ [W] is the noise power at the i-th receiver, which can be expressed as $P_{N,i} = N_0 B_w$, with N_0 [W/Hz] as the noise spectral density and B_w [Hz] is the signal bandwidth.

III. The Genetic Algorithm

In this section, we focus on the specific problem of assigning a certain configuration to each meta-atom, in order to allow the meta-surface to perform beam-steering. With the purpose to provide a detailed explication of the problem we aim to solve, let us consider Fig. 3 (*a*) that is representing a meta-surface constituted by the periodical repetition of the same meta-atoms, whose state can be changed by applying a different voltage value [37]. A corresponding radiation pattern is obtained, based on the specific sequence codes, *i.e.*, sequence of 0 and 1, applied to the meta-surface. The obtained radiation pattern can be estimated by adopting the well-known Huygens equation, as expressed in Eq. (5). Then, we need to tackle the opposite problem, represented in Fig. 3 (*b*). Given a certain radiation pattern (that in this specific work, we infer through the CV support), we need to associate the correct state to each meta-atom, to obtain it. Unfortunately, we cannot apply any formula for that, since the problem is non-linear and there can exist different combinations of meta-atoms states generating the target radiation pattern. It follows that the corresponding problem to be solved can be formulated as follows. Given a certain radiation pattern, represented by the direction of the main beam through the angles characterizing the main beam and its upper mask, we need to assign the correct states to the meta-atoms of our meta-surface, in order to generate the closer radiation pattern configuration.

In order to reconfigure each single meta-atom of the meta-surface presented in the previous section, we have implemented a GA, whose pseudocode is shown in Algorithm 1. This algorithm has been chosen because, thanks to its characteristics, it is suitable to find the configuration of the metasurface that generates the desired radiation pattern (or the closest one). It can evaluate each configuration and, if the upper masks are different, the algorithm generates new configurations through the operations of crossover and mutation.

The specific *population* is constituted by the set of potential configurations with which specific radiation patterns of the RIM are associated. Each configuration is an individual to whom the radiation pattern and upper mask are associated. The upper mask is derived from the 2D radiation pattern (in the plane xy) and it is a matrix comprised of zeros and a single 1 in the position of the main beam. The gene is the state of a single meta-atom. We define an *elite* as the set of the configurations with different upper mask and so with a different radiation pattern.

The searching phase is based on the verification of the presence of the specific configuration in the *elite*. After this phase, if no matching is found, the procedure is applied to the *population* and, if no matching is found, the *cross-over* and *mutation* operations are applied.

In the last stage of the algorithm, the most suitable elements are selected and if their characteristics in terms of upper mask are not responding to the desired ones, the reproduction step is repeated until the fitness function is satisfied or a certain number of generations are met.

The GA capacity of exploring the search space is a desired characteristic for such a type of applications as considered in this work. Different new solutions can be explored through the *cross-over* and *mutation* operations, while keeping the search limited through the correct definition of the fitness function, where specific features of the radiation patterns are defined.

Algorithm 1 Genetic Algorithm (GA)

Require: $f = operating\ frequency$, $m, n = metasurface\ size$, $cross = crossover\ factor$, $mut = mutation\ factor$, $tg = N^{\circ}\ iterations$, $wanted\_Ezz = normalized\ wanted\ radiation\ pattern$, $code\_0$ $code\_1 = S11\ Amplitude\ at\ f\ for\ the\ chosen\ meta\ atom$, $d = distance\ among\ two\ meta\ atoms\ for\ each\ elite\ individual$

Select $chromosomes\ for\ elite$
Compare $upper_{mask} - wanted_{mask}$
if $upper_{mask} == wanted_{mask}$ **then**
  Save $- radiation\ pattern\ \&\ upper_{mask}$
end if
while $no\ optimal\ configuration - elite$ **do**
  Compare $upper_{mask}\ with\ wanted_{mask}$
  if $upper_{mask} == wanted_{mask}$ **then**
    Save $- radiation_{pattern}\ \&\ upper_{mask}$
  else
    Do $crossover$, **Do** $mutation$
    Compute $radiation_{pattern}$, **Check** $sidelobe_{level}$
  end if
end while
if $sidelobe_{level} > max\_threshold$ **then**
  Skip
else
  Extract $upper_{mask}$, **Save** $in\ offspring$, **Save** $in\ elite$
  Compare $upper_{mask} - first_{best-offspring} - wanted\ mask$
end if
if $upper_{mask} == wanted_{mask}$ **then**
  Save $radiation_{pattern}\ \&\ upper_{mask}$
end if
if $individual - not - in - population$ **then**
  Add $individual - in - population$
end if
if $individual - not - in - the - elite$ **then**
  Add $individual - in - elite$
end if

(*a*) (*b*)

Fig. 3. We consider a metasurface, constituted by the periodic repetition of meta-atoms, and we can change the radiation pattern by attributing different states (*i.e.*, 0 and 1) to each meta-atom. (*a*) A certain radiation pattern is obtained starting form a sequence coding of the meta-atoms, that can be computed by applying the Huygens equation. (*b*) If we consider a certain target radiation pattern, we cannot apply an equation to compute it, since the problem is non-linear, and we need to rely on adaptive optimization schemes (*e.g.*, GA) to compute the better radiation pattern configuration.

The GA receives as input the desired upper mask and the algorithm starts the searching phase for the corresponding upper mask and the correspondent radiation pattern. The *fitness function* has been defined in such a way to find the individual, *i.e.* the configuration that is sufficiently "similar" to the desired radiation pattern. In particular, the main beams should have the same directions in terms of azimuth and elevation angles, and the level of side-lobes has to be lower than a certain threshold.

In particular, the different phases associated to GA are as follows:

- STEP 1: Parameter initialization, fitness function evaluation. In this phase, the specific scattered field and corresponding radiation pattern are calculated for different angles, based on the specific fitness function;
- STEP 2: Initialization of the population;
- STEP 3: Selection of the *elite* from the population, based on the best configuration of the population;
- STEP 4: If no optimal configuration is identified, reiterate until a maximum number of iterations is reached;
- STEP 5: Generate offspring with crossover and mutation;
- STEP 6: Compare sidelobe level of the offspring' radiation pattern, compare offspring' upper mask and save the new configuration in the elite.

In order to evaluate how much the direction of the main beam is distant from the desired one, we consider the parameter δ, which is evaluated as the distance between the exact central location of the main beam in the radiation pattern given in input and the one in the radiation pattern corresponding to the coding matrix in output. To have a more precise evaluation of the distance, the position of the two beams has been computed in spherical coordinates, and we express the difference of azimuth (δ_ϕ) and elevation (δ_θ) angles as follows:

$$\delta_\theta = |\theta_p - \theta_e|, \qquad \delta_\phi = |\phi_p - \phi_e|, \tag{8}$$

where the subscript p means the predicted angle, while e the expected angle. If both δ_θ and δ_ϕ are lower than a fixed accuracy value, then the phase configuration that generates that radiation pattern corresponds to the searched one. So, the fitness function is represented by the following equations:

$$\delta_\theta \leq \epsilon_\theta, \qquad \delta_\phi \leq \epsilon_\phi, \tag{9}$$

TABLE I
SIMULATION RESULTS

	μ	σ	MoE_{95}	min_{95}	max_{95}	MoE_{99}	min_{99}	max_{99}
t_{conv}	0.872	0.183	0.047	0.82	0.92	0.061	0.81	0.93

where ϵ_θ and ϵ_ϕ are small values in degrees chosen according to the accuracy required by the system. In this work, we assumed $\epsilon_\theta = 5°$ and $\epsilon_\phi = 10°$.
The time of convergence of the algorithm has been taken into account in Table II. Its mean value is lower than 1 second, so this demonstrates that the algorithm can be adopted in dynamic and real-time scenarios. The convergence time has been evaluated after the training phase (that may take more than 300 s), during which a certain number of configurations to represent the different radiation patterns categories are found. GA complexity is $O(g(nm + nm + n))$ with g as the number of generations, n the population size and m the size of the individuals. Therefore, the complexity is on the order of $O(gnm)$.

When different steering directions are required based on the external conditions, the genetic algorithm searches for the new matrix of the coding schemes, namely the reflective phases of the unit-cells starting from the *elite* and continuing with the *population* and the *cross-over* and *mutation* if no available configuration is found.

If the side-lobe level is lower than the maximum, the *upper mask* of the radiation pattern of the offspring is generated and compared to the *upper mask* of the required radiation pattern. In the meanwhile, the best offspring is saved in the next generation of the *elite*. If the fitness function is satisfied by one of the offspring, the optimal configuration is found; otherwise, the last-discussed procedure is repeated with the updated version of the *elite*. If the maximum number of iterations is reached and no optimum configuration is found, the best phase configuration is chosen among the ones in the *population* and the ones lastly-generated through the *cross-over* and *mutation*.

IV. The Computer Vision approach

In order to provide data to the RIMs, we consider a CV sub-system able to compute the distances between the receiver nodes and the RIMs deployed in a given area. Of course, the CV is able to detect people moving inside the monitoring area, and then estimate the distance from a reference point (*i.e.*, the RIMs).

The CV approach relies on FairMOT, which is a real-time Multi-Object Tracking (MOT) approach [38] that combines object identification and Re-ID in a single network, thus obtaining an high computation efficiency. The object detection pipeline presented in FairMOT is based on the CenterNet architecture, which is an anchor free keypoint estimation approach. As opposed to multi-stage object detection networks, CenterNet is a single-stage end-to-end differentiable model which achieves a good speed-accuracy trade-off, enabling near real-time inference times. In CenterNet, every object is modeled as a point (*e.g.*, the center of mass) and by regression all the other properties, such as bounding box sizes, are derived.

The feature extraction backbone is Deep Layer Aggregation-34 (DLA-34), that is an enhanced version of the ResNet-34 residual neural network. These features are used as an input for both detection and Re-ID branches. If the input image has $H_{img} \times W_{img}$ size, then the object detection branch produces a heatmap of size $C \times H \times W$, where C is the number of channels (*i.e.*, 3 for RGB), $H = H_{img}/4$ and $W = W_{img}/4$.

Applying a SoftMax operation over the entire heatmap, all the peaks are detected and used as object centers. Moreover, the bounding boxes width and height are obtained by regression. The Re-ID branch aims to generate specific features that can distinguish objects. In order to do this, a 2D convolutional layer with 128 kernels is applied on top of the previously extracted features, in order to generate a Re-ID feature vector $E_{x,y} \in R^{128}$ where (x, y) is the location of the object center. Finally, the last step is to keep the link between the detected bounding boxes and the Re-ID features. In order to achieve this, a Kalman Filter is used to predict the object locations in the subsequent frames. The distance between the predicted and detected bounding box locations, and the cosine distance computed on the Re-ID feature vector pairs are then fused in order to establish a if there is a match or not.

Once the network has been trained, the video stream from the cameras is fed as an input and a list of detections including the associated tracking ID and the respective bounding box coordinates is provided as an output. Of course, the bounding box sizes and locations are defined in respect to the 2D image plane (*i.e.*, camera canvas), thus cannot be used directly for the distance calculations. An effective way to relate the transformations between the two planes is to calculate an homography matrix [39], which is the technique we used. For each scenario, a minimum of four key points have been chosen and referenced measuring the exact distance between them in the world and camera plane. More points are recommended in order to take into account any distortions introduced by the camera lens. Assuming that the calibration is precisely performed, the distance in meters between pairs of nodes or between RIM and nodes can be effectively calculated.

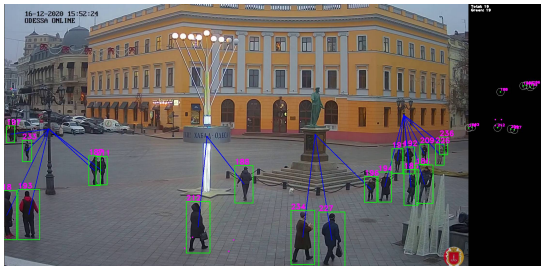

Fig. 4. Example of Odessa video frame processed by the CV, for identification of people within the monitoring area. RIMs are static nodes, that estimate the distance from people (blue links). Reference points are indicated as pink dots on the ground.

In Fig. 4, it is reported how the CV video process performed in a real environment. We consider a real video in Duc de Richelieu Monument in Odessa, Ukraine, recorded by a local webcam [1]. The CV sub-system is able to detect and track several users and estimate the distances between them and the RIMs. Notice that in Fig. 4, RIMs are supposed to be inserted in some elements present in the place *e.g.*, the statue, the street lights, etc., whose positions are *a priori* known.

V. Performance Evaluation

In this section, we evaluate the performance of the overall framework, based on the RIM integration tuned by the means of the GA, receiving the input from a CV sub-system. In order to evaluate the behavior of the global system, we rely on the results of the performance of the individual components. In practice, we need to evaluate separately the performance of the GA algorithm and the CV sub-system, in Subsection V-A and V-B, respectively. Then, the error provided by the CV sub-system is taken as input to the RIM sub-system. This allows evaluating the performance of the overall system in Subsection V-C. The overall methodology adopted, as well as the main performance metrics considered in this paper, are described in Subsection V-D.

A. Genetic Algorithm performance

In order to evaluate the performance of the GA, we have considered two specific RIM structures with size (*i*) $(10\lambda \times 10\lambda)$ and (*ii*) $(5\lambda \times 5\lambda)$ and with four possible states, in order to improve the tunability of the transmitting beams.

The main parameter we are interested to, is the Target Deviation (*i.e.*, $\mathcal{TD}$), namely the difference between the target position and the actual position, and expressed in degree. This will be evaluated for different inputs originated by the CV sub-system. It is worth to highlight that in this work we have evaluated the target deviation for the azimuth angle. In particular, $\mathcal{TD}$ has been computed as the distance between the exact central location of the main beam in the radiation pattern given in input and the one in the radiation pattern corresponding to the coding matrix in output, based on the

[1] Available at the link: https://www.webcamtaxi.com/en/ukraine/odessa/duc-de-richelieu-monument.html.

representation of the beams in spherical coordinates. We have run the simulation for 25 iterations and have considered the average $\mu_{\mathcal{TD}}$, the standard deviation $\sigma_{\mathcal{TD}}$ with the margin of errors (MoE) at 95% and 99%, with minimum and maximum values. The results are shown in Table II. Of course, the results encompass the cases where a corresponding radiation pattern configuration cannot be found, since exact configurations for each possible input cannot be exactly computed. This situation is worst for the smaller configuration $(5\lambda \times 5\lambda)$ and consequently the average error is higher.

B. CV performance

In order to evaluate the performance of the CV sub-system, we considered two aspects *i.e.*, *(i)* the CV capabilities in detecting people collected from a real video, and *(ii)* the estimation error of the distances among people and between people and RIMs. Then, we study the behavior of the joint CV-RIM based system model and its impact in terms of interference reduction.

Object detection networks are commonly evaluated using metrics such as accuracy, precision, recall or F-measure. In this paper we focused on the recall $\mathcal{R}$ [40], because it suited better for this kind of application. The reason is given by the fact that we considered more important to not missing people than obtaining few false positives. The recall metric is defined as:

$$\mathcal{R} = \frac{T_p}{T_p + F_n}, \tag{10}$$

where T_p corresponds to the number of true positives (*i.e.*, a person which has been successfully recognized), and F_n is the total number of false negatives (*i.e.*, people not detected by the system and thus resulting in missing boxes). Our test-bench consisted in a 2 minutes length video clip recorded by a public webcam installed near the Duc de Richelieu Monument in Odessa, Ukraine. The video has a resolution of 1080p (*i.e.* 1920×1080) with a frame rate equal to 25 fps. We sampled an image every 2 s (*i.e.* every 50 frames), thus obtaining a total of 60 input frames. For each frame we manually defined the bounding boxes of all the people present in the scene and used as a basis for comparing the output of the neural network. We obtained a $\mathcal{R} = 94.2\%$, allowing us to adopt this technique as a support for the RIM feeding.

Finally, another critical factor for this system is the precision in computing the distances between the detected nodes. In order to minimize the error, the scene has to be carefully calibrated. As mentioned in Section IV, the calibration consists in picking a minimum of four reference points, and measure the distances between them. Assuming a percentage error equal to Δ in the reference point measurements, we calculated the subsequent error obtained in the CV sub-system distance estimation, over a grid of real points. The Probability Density Function (PDF) of the error (*i.e.*, ε_d [m]) shown in Fig. 5 has been plotted using two different values of Δ (*i.e.* 1% and 3%). The average error obtained is equal to 11 cm with a standard deviation of 5 cm for $\Delta = 1\%$. In the unlikely case of a measurement error of $\Delta = 3\%$, we obtained an average error equal to 34 cm with a standard deviation of 15 cm.

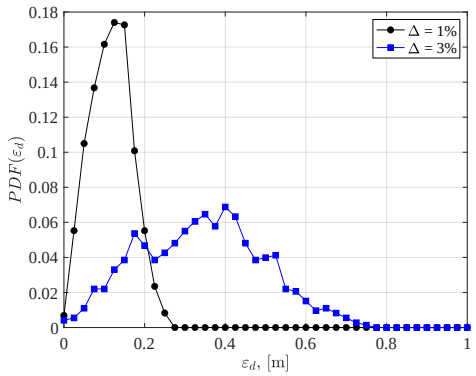

Fig. 5. Probability Density Function of the error in the distance computation, [27].

Based on the user's mobility, the CV approach can introduce an error due to the processing time between the image detection and the user's detection, and then an error of the position estimation. During this interval time, the target user will move a certain distance, based on its velocity and the processing frequency. In Fig. 6, we have reported the errors values, in terms of angles, related to the 90% of cases (the 90-th percentile), due to the processing time. These errors are based on CV processing time and depend on the processing capacity of the system (*i.e.*, frame rate [fps]). We have considered two maximal reference values for the mobility equal to 3 km/h and 6 km/h, as indicated in the 3GPP standard for the typical user's mobility. We assume that users move at a speed uniformly distributed between 0 km/h and the maximal reference value i.e., 3 km/h or 6 km/h. In particular, it is shown that the error can be reduced either if the processing capability of the system is increased (*e.g.*, by introducing Edge Cloud Computing based systems) or by improving the estimation of the reference initial distance, namely by reducing the Δ due to the homography error. It is worth to notice that this error has to be inside the main lobe width, in order to keep the connectivity of the user, that is around 5° for the $(10\lambda \times 10\lambda)$ and 10° in the case of $(5\lambda \times 5\lambda)$.

Still in Fig. 6, we have reported the error based on the estimation of a Radar as reported in [41] and [42]. As we can see, the average error is higher than the 10° required by the system for the 5λ RIM, resulting in an excessive misalignment that does allow to correctly cover the final user considered as target receiver. It is also worth to highlight that the estimation of the angles based on the Radar system does require an ad hoc expensive equipment, that is not the case for the CV, where we aim to exploit the already deployed camera in a Smart City context, resulting in a more affordable and convenient system.

C. Performance System

In order to present the performance of BEST-RIM framework, we need first to introduce the following RIM parameters that will be useful to derive the behavior of our framework:
Directivity, $\mathcal{D}(\theta, \phi)$: it characterizes the amount of energy concentrated in a specific beam direction in respect of an

TABLE II
SIMULATION RESULTS OF THE TARGET DEVIATION FOR THE AZIMUTH ANGLE δ_θ IN CASE OF DIFFERENT RIM SIZES.

RIM size	$\mu_{\mathcal{TD}}$	$\sigma_{\mathcal{TD}}$	MoE_{95}	$\min_{95}$	$\max_{95}$	MoE_{99}	$\min_{99}$	$\max_{99}$
$(10\lambda \times 10\lambda)$	2.95	1.51	0.595	2.245	3.435	0.785	2.06	3.625
$(5\lambda \times 5\lambda)$	7.515	11.865	4.65	3.255	12.55	6.125	1.78	14.03

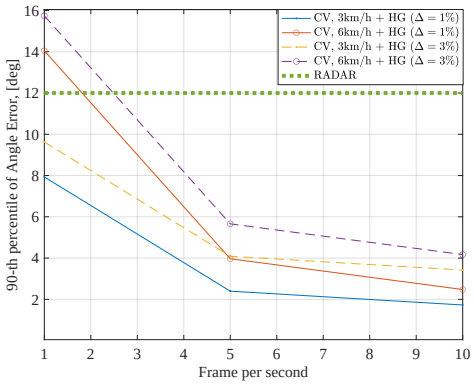

Fig. 6. The 90-th percentile of the angle error vs CV processing time (in terms of frame per second) for 3 km/h and 6 km/h and two homography misalignments $\Delta = 1\%$ and $\Delta = 3\%$.

isotropic scattering considered as reference, and is calculated as:

$$\mathcal{D}(\theta,\phi) = \frac{4\pi\mathcal{P}(\theta,\phi)}{\int_0^{2\pi}\int_0^{\pi}\mathcal{P}(\theta,\phi)\sin\theta d\theta d\phi}, \quad (11)$$

where $\mathcal{P}(\theta,\phi) \propto |E(\theta,\phi)^2|$ is the total power scattered, based on the equation (5) and represents the intensity of the radiation scattered towards a certain direction. The denominator accounts for the total scattered power. If the RIM is fully reflective, the maximum directivity is bounded by $4\pi A/\lambda^2$, where A [m] is the aperture area of the RIM, related to the RIM's size [43]. The directivity will be accounted for the computation of the maximum achievable distance of a receiver.

Target Deviation, $\mathcal{TD}$: it accounts for the difference between the "desired radiation pattern expressed through the angles (θ_r,ϕ_r) and the real one obtained (θ_a,ϕ_a). The target deviation $\mathcal{TD}$ is measured in degree and is computed as:

$$\mathcal{TD} = \sqrt{(\theta_r-\theta_a)^2+(\phi_r-\phi_a)^2} = \mathcal{TD}_{GA}+\mathcal{TD}_{CV}, \quad (12)$$

where $\mathcal{TD}_{GA}$ is the target deviation due to the inaccuracy of the GA and $\mathcal{TD}_{CV}$ is the error committed by the CV approach. Notice that these two sources of errors are independent to each other, and in the worst case, the CV estimation error of the user position will be added to the GA error.

Gain, $\mathcal{G}$: it represents the ratio between the signal intensity G and that one obtained with the radiated power using an isotropic antenna $G_{iso}(\theta,\phi)$ in a certain direction (θ,ϕ), *i.e.*,

$$\mathcal{G}(\theta,\phi) = \frac{G}{G_{iso}(\theta,\phi)}, \quad (13)$$

where $G_{iso}(\theta,\phi) = 1$ for $\theta \in [0,2\pi]$ and $\phi \in [-\pi,\pi]$. Notice that the directivity $\mathcal{D}(\theta,\phi)$ is proportional to the gain $\mathcal{G}(\theta,\phi)$ as follows:

$$\mathcal{D}(\theta,\phi) = \kappa \cdot \mathcal{G}(\theta,\phi), \quad (14)$$

with κ as a proportional constant. If the directivity value is larger than a certain threshold value, the correspondent zone is considered covered by the RIM. Based on the results presented in [43], the directions near to the specular reflection perform better, while the directions approaching the RIM plane present worst performance. In this work, we consider that for a RIM, the maximum directivity is limited to $4\pi A/\lambda^2$, and we derive the gain function by normalizing the directivity in respect to the maximum achievable directivity.

Side Lobe Level: this parameter accounts for the presence of other reflected beams in respect of the main beam. It is mainly due to the finite aperture of the RIM. From a mathematical point of view, we can define it as the ratio of the directivity of the beam closest to the main beam, in respect of the main beam directivity. Of course, low values are desirable for this parameter. We include a maximum value as threshold in the fitness function of the GA, in order to select the phase profiles generating the radiation patterns with low target Side Lobe Level.

Half Power Beam Width: This parameter is measured in degrees and is computed as the square root of the angle at -3 dB of the main lobe. The lower is its value, the better is the tracking accuracy. Anyway, a small value of the Half Power Beam Width, implies a fine tuning technique to be implemented in order to ensure a correct coverage of the mobile user.

In order to evaluate the performance of the overall system, we introduce the Coverage Probability (CP), as the probability that a certain receiver node is reached by a RIM in the mmWave range. In particular, the coverage will be integrate the target deviation, namely the distance between the distance between the target point and the actual point. The target deviation is due to three main factors, the phase inaccuracies, the inaccuracy of the genetic algorithm and the error due to the Computer Vision approach. The first type of inaccuracy is included in the genetic algorithm.

D. Methodology and Results

In this section, we describe the adopted approach for characterizing the RIM and deriving the coverage by integrating the target deviation factor. In this work, we consider a single meta-atom characterized with Floquet boundary conditions simulated in CST [2]. In practice, these boundary conditions correspond to a meta-surface characterized with an infinite number of meta-atoms. The actual RIM is then constituted by a finite number of meta-atoms, where we can apply the global phase and amplitude reflected states, by considering the local quantities.

[2]www.cst.com

We consider a periodic structure and we can use this estimation relying on the slow varying modulation of the RIM as demonstrated in [44]. In order to evaluate the reconfigurable radiation pattern we consider an hybrid approach, partially based on full-wave simulation and an analytic approach, whose derivation is mainly based on the Huygens' principle and the results demonstrated in [43]. The GA has been evaluated with different metasurface sizes, *i.e.*, $(10\lambda\times10\lambda)$ and $(5\lambda\times5\lambda)$. The last two configurations will be considered for the evaluation of the global system.

Complementary Cumulative Distributed Function - CCDF: an important metric for characterizing mmWave performance system is the CP related to the SINR, that characterizes the probability that a certain receiver is "reached" by a RIM. In general, the random variable representing the SINR link is a complex variable encompassing several factors, such as the distances from the interfering transmitter, the channel features accounting for the fading, the beam patterns randomly distributed of the interferers, etc. The CCDF of the SINR for the link from the j-th RIM to the i-th receiver describes the portion of the users achieving a certain threshold x (*i.e.*, $SINR_{ji} \geq x$), averaged over the space and the time. In practice, we compute the probability that the SINR is larger or equal to a certain threshold x, *i.e.*

$$\Pr(SINR_{ji} \geq x) = \Pr\left(\frac{P_i}{\mathcal{I} + P_{N,i}} \geq x\right), \tag{15}$$

where the noise power level at the i-th receiver, equivalent for all the users is expressed as in Eq. (7), and P_i [W] is:

$$P_i = h_{ji} G(\beta,\gamma) G(\theta,\phi) P_j d_{ji}^{-\alpha}, \tag{16}$$

where $G(\beta,\gamma)$ represents the attenuation due to the directivity of the beam with the departure angles (β,γ), while (θ,ϕ) represents the arrival angles. These attenuations will be derived based on the specific characteristics of the RIM, $h_{ji} \sim \exp(1)$, since we are considering Rayleigh channel, P_j [W] is the emitting power of the j-th RIM, which is calculated based on the specific characteristics of the RIM, by considering that the meta-atom is designed for minimizing the power loss of the impinging signal, and $d_{ji}^{-\alpha}$ [m] is the distance between the j-th RIM and the i-th receiver, with two different values of path-loss for characterizing the outdoor path-loss, namely $\alpha = 2.5$ and $\alpha = 3.5$.

Based on the independence of the interference signals and the noise, we can re-write Eq. (15) as:

$$\Pr(SINR_{ji} \geq x) = \Delta(\mathcal{I}(x))\Delta(P_{N,i}(x)), \tag{17}$$

where $\Delta(\cdot)$ represents the Laplacian operator. Finally, since all the parameters are independently distributed, Eq. (17) becomes:

$$\Pr(SINR_{ji} \geq x) = \exp\left[-P_{N,i}(x) - \xi x^{2/\alpha}\right], \tag{18}$$

where, based on the assumption that the angles between a RIM and a receiver are equiprobable and by considering the probability generating functional of PPP [31], the amount ξ can be defined as:

$$\xi = \frac{\mathbb{E}^2[G(\theta,\phi)^{2/\alpha}] C_i \left[\frac{\lambda}{d_{ji}} - \left(\frac{\lambda}{d_{ji}} + \lambda\pi T_H^2\right) e^{-d_{ji}\pi T_H^2}\right]}{sinc(2/\alpha)}, \tag{19}$$

where $\mathbb{E}(\cdot)$ represents the mean value, $C_i = 0$ for traditional communication link, and $C_i = 1$ in case of mmWave. Also, in Eq. (19), T_H [m] is the threshold for communication mode selection, and $sinc(\cdot)$ is the well-known function defined as:

$$sinc(x) = \begin{cases} \frac{\sin(x)}{x} & \text{if } x \neq 0, \\ 1 & otherwise. \end{cases} \tag{20}$$

Power Loss due to Target Deviation: Target deviation causes an additional propagation power loss corresponding to a beam misalignment. Let us denote this additional power loss as ΔL^{mis} and express it in terms of the angle misalignment. Notice that in this work, we only focus on azimuth misalignment $\Delta\phi$ [3], expressed as:

$$\Delta\phi = \phi - arg\min_{\phi}[L(\phi, d, W_\phi)], \tag{21}$$

where W_ϕ is the beamwidth of the radiation pattern reflected by the j-th RIM, and $L(\phi, d_{ji}, W_\phi)$ represents the path loss for the link from the j-th RIM to the i-th receiver. In Eq. (21), $\Delta\phi$ is the difference between the desired angles and the actual ones, whose value is related with the beamwidth. The power loss increases as $\Delta\phi$ increases, and saturates for a large value. Even though the directivity is higher for smaller beamwidths, the precision of beam steering needs to be very high for small values of beamwidths. For small values of beamwidth, a slight misalignment can have a large power loss. Moreover, the power loss is inversely proportional to the beamwidth. In particular, the misalignment additional loss can be expressed as:

$$\Delta L^{mis} \propto \begin{cases} \delta|\Delta\phi|(\frac{1}{W_\phi} - \frac{1}{360^\circ}), & \text{for } W_\phi \leq W_{lim}, \\ 1 & \text{otherwise.} \end{cases} \tag{22}$$

The factor δ is set equal to 3.75 as assumed in [45]. The parameter W_{lim} represents a maximum value of the beamwidth, beyond that the power loss is negligible; in this work we have assumed $W_{lim} = 239.69^\circ$. This target deviation or misalignment implies a reduced power received to the receiver. If the beamwidth is sufficiently large, the additional path loss is not affecting the received power.

In Fig. 7, we show the SINR CCDF versus the SINR, ranging from -20 to 15 dB, in case of different RIM configurations, with $(N = M = 10\lambda)$ meta-atoms, and three different network densities and a path loss coefficient $\alpha = 2.5$ and $\alpha = 3.5$ in Fig. 7 (*a*), (*b*), (*c*) and (*d*), (*e*), (*f*), respectively. For each scenario, the SINR CCDF has been computed for four different cases *i.e.*, (*i*) using only $\mathcal{TD}_{GA}$, (*ii*) using both $\mathcal{TD}_{GA}$ and $\mathcal{TD}_{CV}$, (*iii*) considering no estimation error coming from the GA and the CV approaches, and (*iv*) no RIM deployed in the network. Simulation parameters considered

[3]Similar equations apply for the elevation misalignment.

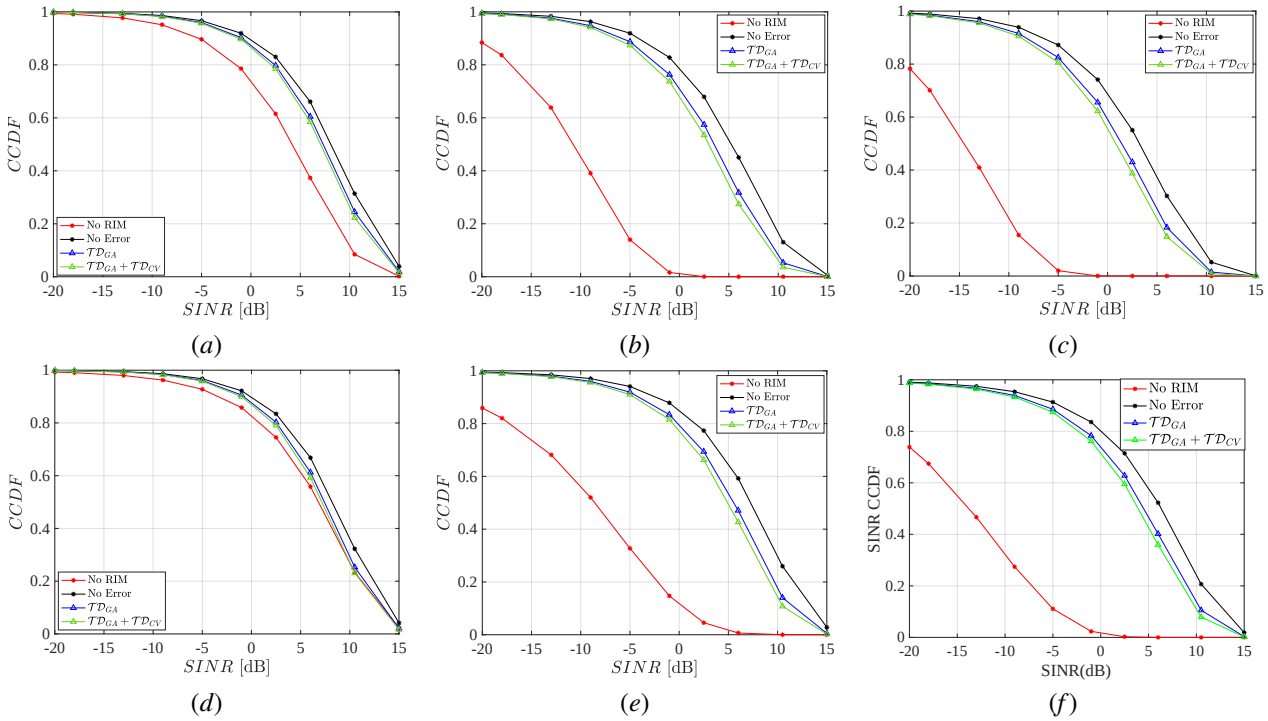

Fig. 7. CP CCDF for a $(10\lambda \times 10\lambda)$ RIM with path loss exponent $\alpha = 2.5$, for (a) sparse, (b) medium density and (c) dense networks, with path loss exponent $\alpha = 3.5$, for (d) sparse, (e) medium density and (f) dense networks.

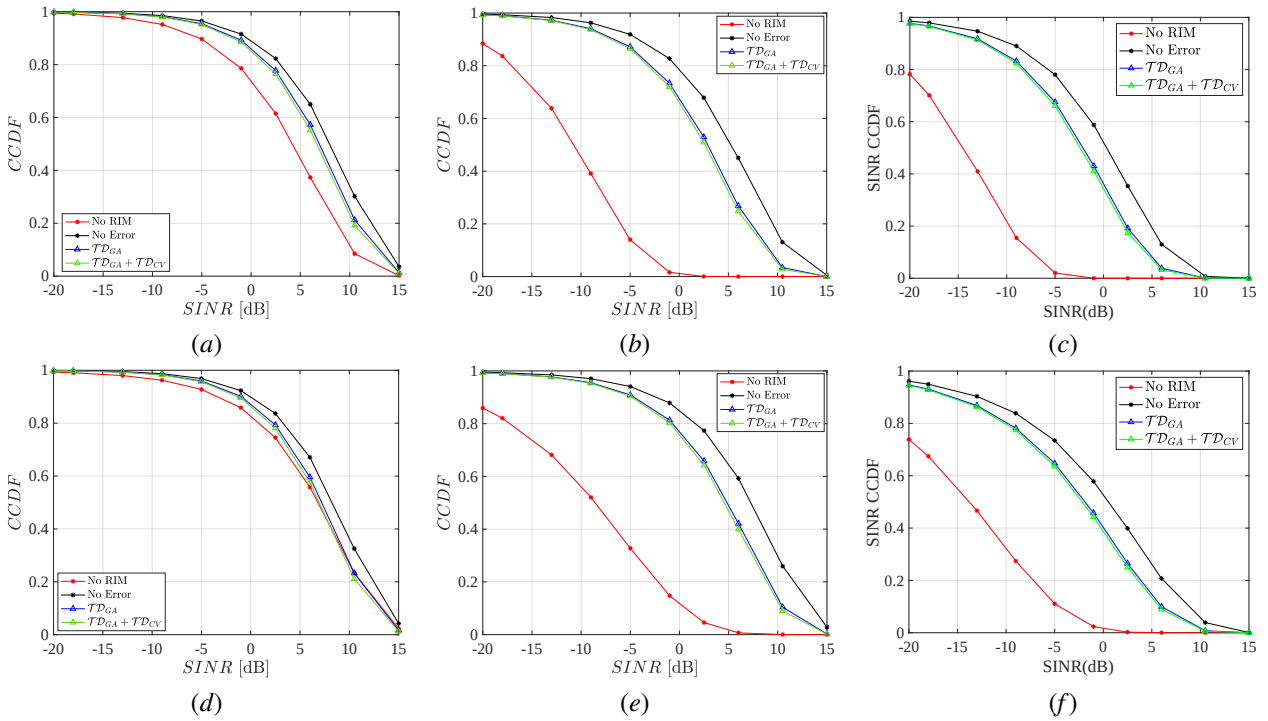

Fig. 8. CP CCDF for a $(5\lambda \times 5\lambda)$ RIM with path loss exponent $\alpha = 2.5$, for (a) sparse, (b) medium density and (c) dense networks, with path loss exponent $\alpha = 3.5$, for (d) sparse, (e) medium density and (f) dense networks.

in this work are reported in Table III. It is worth to notice that the specific value considered for $\mathcal{TD}_{CV}$, is related to a specific scenario, where a user is assumed to move at a maximum speed of 3 km/h, the processing frequency of the video is 5 fps, and the homography error corresponding to $\Delta = 3\%$ is assumed. As explained in the previous section, in order to apply the control capability of the overall architecture based on RIM, presented in this work, we need to keep the average target deviation smaller than the half beam width of the generated radiation pattern, otherwise, as for the case

of the considered baseline, the user cannot be considered correctly covered. We can notice, that for sparse networks the improvement introduced by the communication at mmWave and the beam steering is small, but more evident in a free space condition (*i.e.*, $\alpha = 2.5$). This is mostly due to high distance between the different mmWave transmitters. Indeed, also for the case of omnidirectional antennas, the CP is quite high. In the free-space case, for sparse networks, the possibility to direct the beams allows a better coverage related with low signal losses in the propagation channel. The target deviation due to both the errors based on the CV approach and the GA algorithm has a low impact in all the scenarios considered. A higher dimension of the RIM would allow to obtain a better resolution in terms of beams, but the correspondent beamwidths will be smaller with a reduced tolerance in respect of the target deviation.

More in details, in Fig. 7 (*a-c*) and (*d-f*) we can observe as the environment impacts on the performance of the system. Indeed, in the case of $\alpha = 2.5$ (*i.e.*, propagation conditions close to free space), the amount of interference of each user on the other users is higher, due to the higher distance of the propagation signal. This is more visible for the case of "No RIM", namely in absence of directional beams (*red curves*). On the other side, in Fig. 7 (*a-c*) and (*d-f*), the RIM-based solution is slightly affected by the environment. Interestingly, we can observe as the distance between the "No Error" approach, that represents our ideal solution (*black curves*), is very similar in the three user densities and for $\alpha = 2.5$ and $\alpha = 3.5$. In these cases, we can also note that curves with GA and CV errors (*i.e.*, *blue* and *green curves*) are quite close to the ideal case, thus the error inducted on the GA and CV can be considered independent of the channel state conditions and the user density.

The impact of directionality is more evident with the increasing of the user density, as we can observe in Fig. 7 (*b*) and (*e*), and by comparing them to Fig. 7 (*c*) and (*f*), respectively. By considering Fig. 7 (*b*) and (*c*), we can observe as the directional approaches are slightly impacted by the increasing of user density. This result is confirmed by the comparison of Fig. 7 (*d*), (*e*) and (*f*), where the user density impact is a bit lower, but still there exists a degradation of performance at higher densities.

In Fig. 8, we show the performance of the RIM-based structure, with smaller meta-surface, implying a reduced number of potential configurations and a reduced directional capability, since the number of potential configurations is smaller. Anyway, it is interesting the comparison of the results with the same parameter configurations, for the two different RIM sizes. Indeed, the behavior for smaller densities (see Fig. 7 (*a*)/(*d*) and Fig. 8(*a*)/(*d*) respectively, and Fig. 7(*b*)/(*e*) and Fig. 8(*b*)/(*e*) respectively) shows the size of the RIM has a very low impact. Only when the user density is very high (see Fig. 7 (*c*)/(*f*) and Fig. 8 (*c*)/(*f*)), the higher directional capability plays a more important role. These results are very important to designing effective RIM-based structures and minimizing the cost. Indeed, higher sizes can provide better directionality, but the number of components for RIM (*e.g.*, the pin diodes) explodes, with an important impact on the cost. A provision of traffic load in terms of users can help to decide the most suitable RIM size.

TABLE III
SIMULATION PARAMETERS

Band frequency, f	28 GHz
Path Loss exponent, α	2.5, 3.5
T_H	150 m
RIM size	$(10\lambda \times 10\lambda)$ $(5\lambda \times 5\lambda)$
SNR	10 dB
$P_{N,i}$	10^{-1} dB
$\mathcal{TD}_{CV}$	6°
$\mathcal{TD}_{GA}$ for $(10\lambda \times 10\lambda)$	3°
$\mathcal{TD}_{GA}$ for $(5\lambda \times 5\lambda)$	7.5°
Beamwidth for $(10\lambda \times 10\lambda)$	10°
Beamwidth for $(5\lambda \times 5\lambda)$	20°
Communication Selection, C	1
$\max(\mathcal{D})$ for $(10\lambda \times 10\lambda)$	25 dB
$\max(\mathcal{D})$ for $(5\lambda \times 5\lambda)$	20 dB
User density [m^{-2}]	
sparse network	$\frac{2}{(\pi 500^2)}$
medium network	$\frac{50}{(\pi 500^2)}$
dense network	$\frac{100}{(\pi 500^2)}$

VI. CONCLUSION AND DISCUSSION

In this work we have proposed the BEST-RIM framework, considering a RIM whose working frequency is 28 GHz. The coding schemes of the RIM structure are dynamically computed by means of a PU running a GA whose performance in terms of accuracy and average error have been assessed with different RIM configurations. The GA is fed with the inputs deriving from a CV approach that has been evaluated in an outdoor pedestrian environment. The whole communication system has been evaluated by considering the CP derived on the SINR basis. Results show that for a $(10\lambda \times 10\lambda)$ RIM, the integration of the beam steering capability allows to improve the performance in terms of coverage in respect of an omnidirectional isotropic ideal antenna, above all for dense scenarios.

Results obtained in this work show the high potentiality of mmWave 5G communication relying on RIM configurations. Anyway, this analysis allowed to understand that such a kind of framework is a complex system, where each layer has to be carefully implemented since it impacts on the other components of the system. In a mobile scenario as we have considered in this work, the beam misalignment can be also generated by the delay of the PU and the delay associated with the FPGA that need to translate the values voltages for making them corresponding to the desired reconfigured radiation pattern. All these factors need to be included in the evaluation of the system, in order to better assess the performance of such a kind of systems.

ACKNOWLEDGMENT

This work was partially supported by ETHICAM Project, CPER DATA and ELSAT projects.

REFERENCES

[1] A. Ghosh, T. A. Thomas, M. C. Cudak, R. Ratasuk, P. Moorut, F. W. Vook, T. S. Rappaport, G. R. MacCartney, S. Sun, and S. Nie, "Millimeter-Wave Enhanced Local Area Systems: A High-Data-Rate Approach for Future Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1152–1163, 2014.

[2] S. Buzzi, C. D'Andrea, M. Fresia, and X. Wu, "Beam Alignment in mmWave User-Centric Cell-Free Massive MIMO Systems," 2021. [Online]. Available: https://arxiv.org/abs/2106.13538

[3] "White paper on 5G channel model for bands up to 100 GHz," NTT DOCOMO, Tech. Rep., 2016.[Online]. Available: http://www.5gworkshops.com/5GCM.html (Last accessed: Nov. 24, 2022), Tech. Rep., 2016.

[4] M. Xiao, S. Mumtaz, Y. Huang, L. Dai, Y. Li, M. Matthaiou, G. K. Karagiannidis, E. Bjornson, K. Yang, C.-L. I, and A. Ghosh, "Millimeter Wave Communications for Future Mobile Networks," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 9, pp. 1909–1935, 2017.

[5] A. Li and C. Masouros, "Hybrid Analog-Digital Millimeter-Wave MU-MIMO Transmission With Virtual Path Selection," *IEEE Communications Letters*, vol. 21, no. 2, pp. 438–441, 2017.

[6] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-wave massive MIMO: the next wireless revolution?" *IEEE Communications Magazine*, vol. 52, no. 9, pp. 56–62, 2014.

[7] S. Li, B. Duo, X. Yuan, Y.-C. Liang, and M. Di Renzo, "Reconfigurable Intelligent Surface Assisted UAV Communication: Joint Trajectory Design and Passive Beamforming," 2019. [Online]. Available: https://arxiv.org/abs/1908.04082

[8] N. S. Perovic, M. D. Renzo, and M. F. Flanagan, "Channel Capacity Optimization Using Reconfigurable Intelligent Surfaces in Indoor mmWave Environments," in *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020, pp. 1–7.

[9] H. Gacanin and M. Di Renzo, "Wireless 2.0: Toward an Intelligent Radio Environment Empowered by Reconfigurable Meta-Surfaces and Artificial Intelligence," *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 74–82, 2020.

[10] e. a. T. Cui, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light Sci Appl*, 2014.

[11] C. Liaskos, S. Nie, A. Tsioliaridou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "A New Wireless Communication Paradigm through Software-Controlled Metasurfaces," *IEEE Communications Magazine*, vol. 56, no. 9, pp. 162–169, 2018.

[12] X. Yu, D. Xu, D. W. K. Ng, and R. Schober, "IRS-Assisted Green Communication Systems: Provable Convergence and Robust Optimization," *IEEE Transactions on Communications*, vol. 69, no. 9, pp. 6313–6329, 2021.

[13] X. Tan, Z. Sun, D. Koutsonikolas, and J. M. Jornet, "Enabling Indoor Mobile Millimeter-wave Networks Based on Smart Reflect-arrays," in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*, 2018, pp. 270–278.

[14] S. Zhang, M. Li, M. Jian, Y. Zhao, and F. Gao, "AIRIS: Artificial intelligence enhanced signal processing in reconfigurable intelligent surface communications," *China Communications*, vol. 18, no. 7, pp. 158–171, 2021.

[15] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, "Smart Radio Environments Empowered by Reconfigurable Intelligent Surfaces: How It Works, State of Research, and The Road Ahead," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[16] Q. Wu and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Network via Joint Active and Passive Beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.

[17] ——, "Beamforming Optimization for Intelligent Reflecting Surface with Discrete Phase Shifts," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019, pp. 7830–7833.

[18] P. Mursia, V. Sciancalepore, A. Garcia-Saavedra, L. Cottatellucci, X. Costa-Perez, and D. Gesbert, "RISMA: Reconfigurable intelligent surfaces enabling beamforming for IoT massive access," *IEEE Journal on Selected Areas in Communications, Special Issue on Massive Access for 5G and Beyond, 24 August 2020*, 2020.

[19] M.-M. Zhao, Q. Wu, M.-J. Zhao, and R. Zhang, "Intelligent Reflecting Surface Enhanced Wireless Networks: Two-Timescale Beamforming Optimization," *IEEE Transactions on Wireless Communications*, vol. 20, no. 1, pp. 2–17, 2021.

[20] A. Abrardo, D. Dardari, and M. D. Renzo, "Intelligent Reflecting Surfaces: Sum-Rate Optimization Based on Statistical Position Information," *IEEE Transactions on Communications*, vol. 69, pp. 7121–7136, 2021.

[21] K. Zhi, C. Pan, H. Ren, and K. Wang, "Statistical CSI-Based Design for Reconfigurable Intelligent Surface-Aided Massive MIMO Systems With Direct Links," *IEEE Wireless Communications Letters*, vol. 10, pp. 1128–1132, 2021.

[22] M. Jeong, B. C. Ko, and J.-Y. Nam, "Early Detection of Sudden Pedestrian Crossing for Safe Driving During Summer Nights," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 27, no. 6, pp. 1368–1380, 2017.

[23] Z. Yuan, T. Azzino, Y. Hao, Y. Lyu, H. Pei, A. Boldini, M. Mezzavilla, M. Beheshti, M. Porfiri, T. E. Hudson, W. Seiple, Y. Fang, S. Rangan, Y. Wang, and J.-R. Rizzo, "Network-Aware 5G Edge Computing for Object Detection: Augmenting Wearables to "See" More, Farther and Faster," *IEEE Access*, vol. 10, pp. 29 612–29 632, 2022. [Online]. Available: https://doi.org/10.1109%2Faccess.2022.3157876

[24] T. Nishio, Y. Koda, J. Park, M. Bennis, and K. Doppler, "When Wireless Communications Meet Computer Vision in Beyond 5G," *IEEE Communications Standards Magazine*, vol. 5, no. 2, pp. 76–83, 2021.

[25] L. Wei, Y. Guan, L. Chen, and L. Sun, "Millimeter-Wave Radar and Machine Vision-Based Lane Recognition," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 32, 2018.

[26] J. Duque Domingo, J. Gómez-García-Bermejo, E. Zalama, C. Cerrada, and E. Valero, "Integration of Computer Vision and Wireless Networks to Provide Indoor Positioning," *Sensors*, vol. 19, no. 24, 2019. [Online]. Available: https://www.mdpi.com/1424-8220/19/24/5495

[27] V. Loscrí, A. M. Vegni, E. Innocenti, R. Giuliano, and F. Mazzenga, "A joint Computer Vision and Reconfigurable Intelligent Meta-surface Approach for Interference Reduction in Beyond 5G Networks," in *2021 IEEE 22nd International Conference on High Performance Switching and Routing (HPSR)*, 2021, pp. 1–6.

[28] T. Nitsche, A. B. Flores, E. W. Knightly, and J. Widmer, "Steering with eyes closed: mm-Wave beam steering without in-band measurement," in *2015 IEEE Conference on Computer Communications (INFOCOM)*, 2015, pp. 2416–2424.

[29] A. Ali, N. González-Prelcic, and R. W. Heath, "Millimeter Wave Beam-Selection Using Out-of-Band Spatial Information," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 1038–1052, 2018.

[30] H.-W. Liu, T.-X. Zheng, Y. Wen, C. Feng, and H.-M. Wang, "Performance Analysis of Uplink mmWave Communications in C-V2X Networks," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, 2020, pp. 1–6.

[31] F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks*, 2009, vol. 2.

[32] J. Huang, C.-X. Wang, Y. Sun, R. Feng, J. Huang, B. Guo, Z. Zhong, and T. J. Cui, "Reconfigurable Intelligent Surfaces: Channel characterization and modeling," 2022. [Online]. Available: https://arxiv.org/abs/2206.02308

[33] A. K. Gupta, J. G. Andrews, and R. W. Heath, "On the feasibility of sharing spectrum licenses in mmwave cellular systems," *IEEE Transactions on Communications*, vol. 64, no. 9, pp. 3981–3995, 2016.

[34] H. Yang, X. Cao, F. Yang, J. Gao, S. Xu, M. Li, X. Chen, Y. Zhao, Y. Zheng, and S. Li, "A programmable metasurface with dynamic polarization, scattering and focusing control," *Scientific reports*, vol. 6, no. 1, pp. 1–11, 2016.

[35] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*. Wiley New York, 1995.

[36] J.-F. Coeurjolly, J. Moller, and R. Waagepetersen, "A tutorial on Palm distributions for spatial point processes," 2015. [Online]. Available: https://arxiv.org/abs/1512.05871

[37] C. Rizza, V. Loscri, and M. O. Parchin, "Real-Time Beam steering in mmWave with Reconfigurable Intelligent Meta-surfaces," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[38] Y. Zhang, C. Wang, X. Wang, W. Zeng, and W. Liu, "FairMOT: On the fairness of detection and re-identification in multiple object tracking," *International Journal of Computer Vision*, vol. 129, no. 11, pp. 3069–3087, sep 2021. [Online]. Available: https://doi.org/10.1007%2Fs11263-021-01513-4

[39] D. Kriegman, "Homography estimation," in *Computer Vision ICSE252A, Winter2007*. [Online]. Available: Available at http://cseweb.ucsd.edu/classes/wi07/cse252a/homography_estimation/homography_estimation.pdf

[40] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006,

rOC Analysis in Pattern Recognition. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016786550500303X

[41] K. Gupta, S. Joshi, M. B. Srinivas, S. Boppu, M. S. Manikandan, and L. R. Cenkeramaddi, "Localization of Multi-Class On-Road and Aerial Targets Using mmWave FMCW Radar," *Electronics*, vol. 10, no. 23, 2021. [Online]. Available: https://www.mdpi.com/2079-9292/10/23/2905

[42] L. R. Cenkeramaddi, P. K. Rai, A. Dayal, J. Bhatia, A. Pandya, J. Soumya, A. Kumar, and A. Jha, "A Novel Angle Estimation for mmWave FMCW Radars Using Machine Learning," *IEEE Sensors Journal*, vol. 21, no. 8, pp. 9833–9843, 2021.

[43] H. Taghvaee, S. Abadal, A. Pitilakis, O. Tsilipakos, A. C. Tasolamprou, C. Liaskos, M. Kafesaki, N. V. Kantartzis, A. Cabellos-Aparicio, and E. Alarcon, "Scalability Analysis of Programmable Metasurfaces for Beam Steering," *IEEE Access*, vol. 8, pp. 105 320–105 334, 2020.

[44] F. Liu, O. Tsilipakos, A. Pitilakis, A. C. Tasolamprou, M. S. Mirmoosa, N. V. Kantartzis, D.-H. Kwon, M. Kafesaki, C. M. Soukoulis, and S. A. Tretyakov, "Intelligent Metasurfaces with Continuously Tunable Local Surface Impedance for Multiple Reconfigurable Functions," *Physical Review Applied*, vol. 11, no. 4, apr 2019. [Online]. Available: https://doi.org/10.1103%2Fphysrevapplied.11.044024

[45] J. Lee, M.-D. Kim, J.-J. Park, and Y. J. Chong, "Field-Measurement-Based Received Power Analysis for Directional Beamforming Millimeter-Wave Systems: Effects of Beamwidth and Beam Misalignment," *ETRI Journal*, vol. 40, no. 1, pp. 26–38, 2018. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.4218/etrij.2017-0188