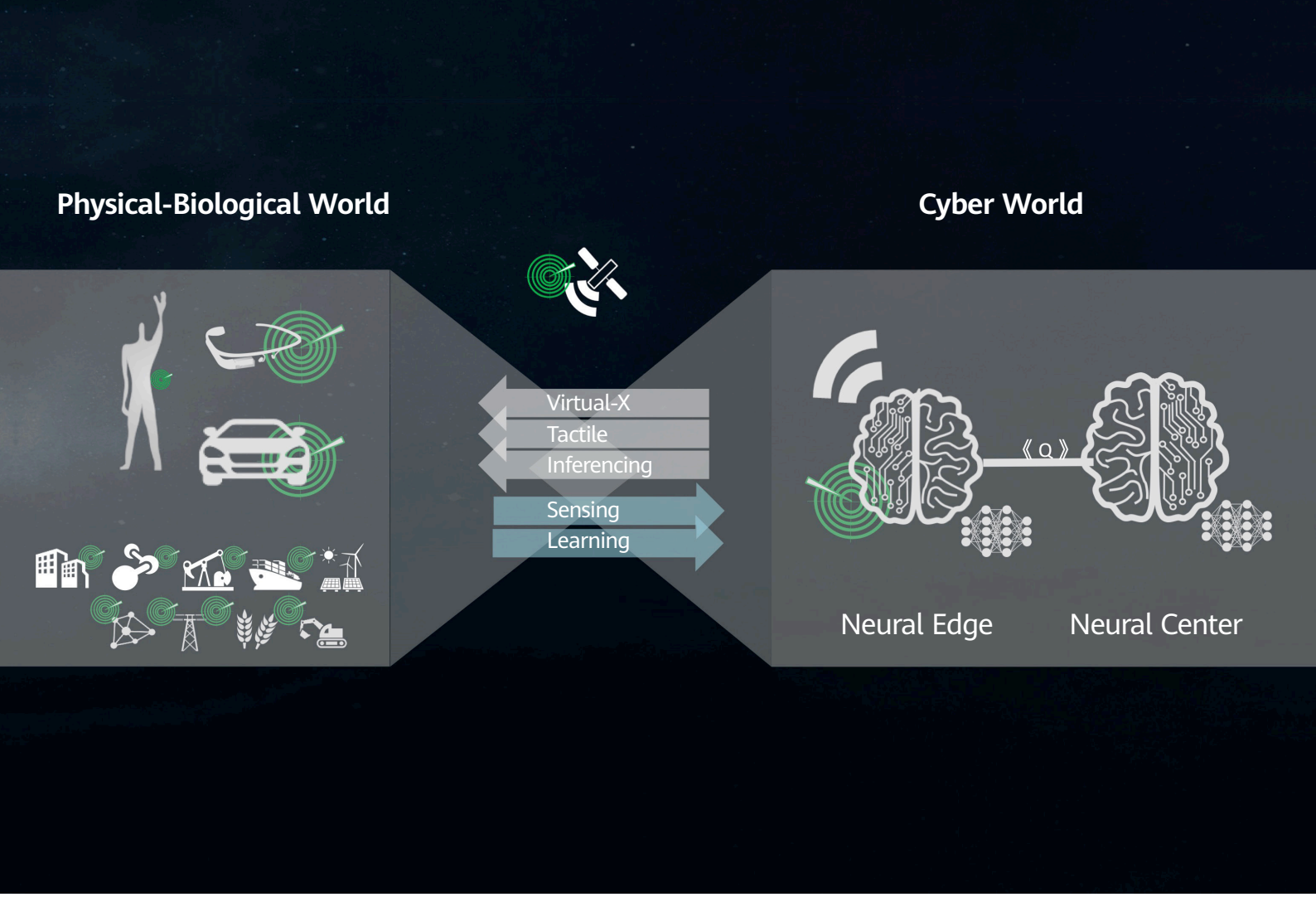



# Integrated Sensing and Communication (ISAC) — From Concept to Practice

Alireza Bayesteh [1], Jia He [2], Yan Chen [1], Peiying Zhu [1], Jianglei Ma [1], Ahmed Wagdy Shaban [1], Ziming Yu [2], Yunhao Zhang [2], Zhi Zhou [2], Guangjian Wang [2]

[1] Ottawa Wireless Advanced System Competency Centre

[2] Wireless Technology Lab

**Abstract**

6G will serve as a distributed neural network for the future Intelligence of Everything. Network Sensing and Native AI will become two new usage scenarios in the era of connected intelligence. 6G will integrate sensing with communication in a single system. Radio waves can be exploited to "see" the physical world and make a digital twin in the cyber world. This article introduces the concept of integrated sensing and communication (ISAC) and typical use cases, and provides two case studies of how to use 6G ISAC to improve localization accuracy and perform millimeter level imaging using future portable devices. The research challenges to implementing ISAC in practice are discussed.

**Keywords**

integrated sensing and communication (ISAC), localization, THz imaging, sensing accuracy, sensing resolution, prototype

# 1 Introduction

In 6G mobile communication systems, the use of higher frequency bands (from mmWave up to THz), wider bandwidth, and massive antenna arrays will enable high-accuracy and high-resolution sensing, which can help implement the integration of wireless signal sensing and communication (ISAC) in a single system for their mutual benefit. On the one hand, the entire communications network can serve as a sensor. The radio signals transmitted and received by network elements and the radio wave transmissions, reflections, and scattering can be used to sense and better understand the physical world. The capabilities to obtain range, velocity, and angle information from the radio signals can provide a broad range of new services, such as high accuracy localization, gesture capturing and activity recognition, passive object detection and tracking, as well as imaging and environment reconstruction [1]. This is called "network as a sensor". On the other hand, the capabilities of high-accuracy localization, imaging, and environment reconstruction obtained from sensing can improve communication performance — for example, more accurate beamforming, faster beam failure recovery, and less overhead when tracking the channel state information (CSI) [2–3]. This is called "sensing-assisted communication". Moreover, sensing is a "new channel" that observes, samples, and links the physical and biological world to the cyber world. Real-time sensing is therefore essential to make the concept of the digital twin — a true and real-time replica of the physical world — a reality in the future.

3GPP has initiated some preliminary study on use cases and potential ISAC requirements using the air interface of 5G advanced [1]. 6G ISAC systems will, however, be further optimized, fully integrated, and will not be constrained by the limitations of the current 5G system. The sensing use cases offered by future 6G ISAC systems will most likely include ultra-high accuracy localization and tracking, simultaneous imaging, mapping, and localization,

**Table 1** ISAC use cases as new services in 6G according to different categories

| Use Case Category / Application Category | High-Accuracy Localization and Tracking | Simultaneous Imaging, Mapping, and Localization | Augmented Human Sense | Gesture and Activity Recognition |
|---|---|---|---|---|
| **Vertical Industry**<br>Intelligent healthcare<br>Intelligent transportation<br>Intelligent factory/ manufacturing<br>Smart agriculture | • Surgery with cooperative robots<br>• Docking drone on a moving vehicle<br>• Device/module placement and installation<br>• Livestock movement and animal migration monitoring | • Sensing glasses with ultra-high resolution imagery<br>• 3D road environment mapping<br>• Warehouse robotics automation system<br>• Crop production and crop physiology | • Tele-surgery<br>• Pollution and air quality detection<br>• Automatic flaw detection on products<br>• Intelligent crop monitoring for nutrients, water stress and disease | • Gesture-controlled smart operation theater<br>• In-cabin monitoring and contactless control<br>• Contactless control for intelligent manufacturing system<br>• Gesture-based robots and machinery control for precision agriculture |
| **Consumer**<br>Smart home & entertainment<br>Smart mobile devices | • Collaborative robots for household chores<br>• Precise localization of small objects (tag or active objects) using mobile phones | • Close-in scene and object imaging | • Imaging of water pipes behind walls<br>• Calories count<br>• Contaminated ingredients detection | • Virtual piano<br>• Touchless home appliances<br>• Contactless control on intelligent screen |
| **Public Service**<br>Smart city<br>Smart environment<br>Smart security and public safety | • Drone as robotic waiter<br>• Hydrological monitoring {e.g., precipitation, water flow/level}<br>• Crowd management and emergency evacuation for major events | • Wireless SLAM<br>• Drone base stations swarm SAR imaging<br>• In-car sensing for driver and passenger monitoring | • Crack detections in buildings, bridges and man-made structures<br>• Fine particulate matter detection (PM10, PM2.5)<br>• Explosive detection and gas sensing<br>• Security scans on packages | • Gesture-based appliances for enhanced accessibility for seniors and differently abled people<br>• Panic and terrifying emotion recognition |

augmented human sensing, gesture and activity recognition, as illustrated in Table 1 [1]. The use cases and performance requirements will be further discussed in Section 2.

The integration of sensing and communication functions can happen at three different levels, from loosely coupled to fully integrated. At the lowest integration level, sensing and communication capabilities can co-exist on hardware by sharing the spectrum, which is more efficient than dedicated spectrum usage. Sensing can benefit from the economies of scale in the mobile communication network, where shared hardware will be cost effective and eases deployment and maintenance issues. The second level of integration calls for the integration of waveform and signal processing where the time, frequency, and spatial domain processing techniques have a common objective and can be combined to serve both sensing and communication functions. A fully integrated system with cross-layer, cross-module, and cross-node information sharing is expected to significantly enhance the mutual performance of both sensing and communication, as well as reduce the overall cost, size and power consumption of the network system.

In addition to the wider spectrum and the larger number of antennas, the sensing functionality and performance will be further enabled by other technology innovations such as the larger scale of cooperation between base stations and user equipment (UE), joint design of communication and sensing waveforms, advanced techniques for interference cancellation, and the native AI capability to better deal with the sensed data.

Next, we will discuss typical ISAC use cases and then elaborate on examples of ISAC application in enhanced localization and millimeter level resolution. Design challenges will be discussed thereafter, followed by the conclusions.

# 2 ISAC Use Cases

## 2.1 Overview

Wireless sensing has long been a separate technology developed in parallel with the mobile communication systems. Positioning is the only sensing service that mobile communication systems (until 5G) could offer. General sensing rather than positioning will become a new function integrated into the 6G mobile communication system. This capability will open up brand new services for 6G. These services are currently provided by various dedicated sensing equipment, such as radar, light detection and ranging (LIDAR), and professional CT and MRI equipment.

The ISAC capability will thus enable many new services that mobile communication system operators can offer. These include very high accuracy positioning, localization and tracking, imaging for biomedical and security applications, simultaneous localization and mapping to automatically construct maps of complex indoor or outdoor environments, pollution or natural disaster monitoring, gesture and activity recognition, flaw and material detection and many other services. These services will in turn enable application scenarios in all kinds of business for future consumers and vertical industries. The potential new services that could be supported by future ISAC systems are listed in Table 1. In the table, the use cases are categorized into four functional categories across different applications/industries (vertical industry, consumer and public services):

- High-accuracy localization and tracking
- Simultaneous imaging, mapping and localization
- Augmented human sensing
- Gesture and activity recognition

It is also worth mentioning that, in addition to the preceding services, sensing can also be used to assist communications and positioning, more details of which can be found in Section 5.4.

## 2.2 High-Accuracy Localization and Tracking

Low-latency high-accuracy localization and tracking enable meaningful association between cyber information and the locations of physical entities in multiple scenarios from factories to warehouses, hospitals to retail shops, and agriculture to mining.

The 6G network will provide services for both device-based and device-free object localization. For 6G device-based localization, the target is a connected device in the network, and the location information is derived from the reference signals or measurement feedback from the

device. Localization for 6G device-free objects, on the other hand, does not need the object to be a connected device in the network. The estimation of delay, Doppler, and angle spectrum information (corresponding to the distance, velocity, and angle of the objects) are obtained from the scattered and reflected wireless signals either through monostatic sensing (receiver is the same as transmitter) or bistatic sensing (receiver is another node or device in the network). By processing these wireless signals further, the locations, orientations, velocity, and other geometric information of the objects in a physical 3D space can be extracted. With higher bandwidth and increased antenna aperture, the 6G ISAC system can have strong capabilities to separate multipaths, through which better localization and tracking performance can be achieved, and the localization accuracy for outdoor use cases can be up to the centimeter level.

Having high-accuracy relative localization is important when two or more entities exist and they are approaching one another, or the entities have coordinated moving direction and speed. In automatic warehousing applications, centimeter-level accuracy enables device-level placement, and the near-millimeter-level accuracy can further enable module-level installation and placement in tight spaces, allowing for efficient storage of components that have a small form factor. Relative localization is necessary as a viable alternative for close-in maneuvering owing to the fact that complexity, physical limitations, and external infrastructure are mission critical for each robot to accurately determine its location in relation to a common datum. An example will be a drone docking onto a moving vehicle with an extremely small margin for landing, due to the limited area of the moving vehicle's cargo platform.

Future ISAC systems that are empowered by native AI can provide semantic localization capability with context awareness. To support future smart home/shopping mall/restaurant/hotel, and automatic factory applications, objects and parts need to have dispatchable localization information such as shelf level, seat number, table number, etc. In a restaurant, robotic waiters, which have semantic localization capability, can accurately deliver food to guests and even go a step further and set different level of protections according to different task specifics, e.g., fragile and rigid objects can be treated with different levels of location and velocity accuracy during transportation.

## 2.3 Imaging, Mapping, and Environment Reconstruction

In simultaneous imaging, mapping, and localization, the sensing capabilities from these three perspectives are mutually enhanced. Particularly, the imaging function is used to capture the images of the surrounding environment, and the localization function is used to obtain the locations of surrounding objects. These images and/or locations are then used by the mapping function to construct a map. The mapping function helps the localization function improve the inference of locations. ISAC will leverage on advanced algorithms, edge computing, and AI to produce super-resolution and highly recognizable images and maps in which the vast network of objects, including vehicles and base stations, act as sensors to provide a remarkably extended imaging area. Moreover, performance will significantly improve due to the ease of fusing results that are shared with cloud-based services across the entire network.

6G-based super-resolution and high accuracy sensing applications open up a range of possibilities in 3D indoor imaging and mapping, which in turn enable various applications, such as scene reconstruction, spatial localization, and navigation for indoor scenarios, and help provide the most up-to-date knowledge of an environment for networks and devices. The accurate mapping information can then be applied to determine the multipath reflection points. Owing to the fact that scattered signals bounce multiple times where the LOS surfaces act as mirrors, compensated images of NLOS objects can be reconstructed by applying mirroring techniques. Once the environment is reconstructed, the next step will be the localization and imaging of the NLOS targets. Target locations can then be detected with good accuracy when prior information regarding the scene's geometry is known.

In an outdoor imaging and mapping scenario involving a mobile vehicle, its sensors usually have a restricted view and limited coverage due to weather, obstacles, and the sensors' power control. That said, nearby stationary base stations may have a greater field of view, longer sensing distance, and higher resolution because they collect and use their own sensing data or sensing data of UE. Therefore, mobile vehicles can achieve higher levels of autonomy by utilizing the maps reconstructed by the base stations to determine

their next move. Moreover, the sensing resolution and accuracy performance will significantly improve due to the fusion of imaging results across the network. The densely distributed base stations in an urban area and ISAC make environmental reconstruction and 3D localization possible, which in turn form the virtual city. The reconstructed map used for smart traffic control scenarios, such as traffic flow monitoring, queue detection, and accident detection, are an important use case in the dynamic virtual city.

## 2.4 Augmented Human Senses

Technology progress makes augmented human sensing a reality. Augmented human sensing aims to provide a safe, high-precision, low-power, sensing and imaging capability that exceeds human abilities, by means of a portable terminal (e.g., 6G-enabled mobile phones, wearables, or medical equipment implanted beneath human skin), to sense the surrounding environments. With the help of scientific and technological advancement, augmented human sensing can be achieved to facilitate information collection and integrate the maximum number of environmental messages into the 6G network.

In the 6G network, high-resolution imaging and detection sensing techniques will open the door for numerous applications, such as remote surgery, cancer diagnostics, detection of slits on products, and sink water-leakage detection. A surgeon may be able to conduct surgery at a different location through the help of an ultra-high-resolution imaging monitor system and remote operation platform system. In addition, intelligent factories will leverage these superior sensing solutions to implement contactless ultra-high-precision detection, tracking, and quality control, where millimeter-level radial-range resolution and ultra-high cross-range resolution based on higher bandwidth and increased antenna array aperture, respectively, are required. 6G communication technologies, with high THz frequency and corresponding short wavelength that is less than 1 mm can increase the bandwidth and decrease the array size, so that these augmented human sensing functions can be integrated or installed in portable devices.

While ultra-high-resolution scenarios require higher bandwidth and increased antenna aperture, another application of "seeing beyond the eye" that can sense the changes beneath the skin, behind occlusion, or in darkness, poses different requirements. 6G radio wave (up to THz) based sensing can achieve the NLOS imaging ability, where technology for detecting hidden objects can be equipped on portable devices that have a powerful imaging capability. As such, mobile phones can be used to detect pipelines behind walls or perform security scans on packages by utilizing the penetration characteristics of electromagnetic waves. Moreover, 6G ISAC can enable atraumatic medical detection which plays an important role in eHealth procedures such as diagnosis, monitoring, and treatments. It provides ultra-high reliability and accuracy and does not harm human bodies.

Spectrogram recognition is another interesting application that could be supported by a 6G ISAC system. It can identify targets through spectrogram sensing of their electromagnetic or photonic characteristics. This includes the analysis of absorption, reflectivity, and permittivity parameters, which helps distinguish the type and quality of materials. Pollution and product quality management are some of the prospective applications of this technology. Spectrogram recognition can also be used in food sensing applications to detect the food type and ingredients through the transmission and reflection of THz signals. This technology will help identify different types of food, calorie content, presence of contaminated ingredients, etc.

## 2.5 Posture and Gesture Recognition

Device-free gesture and posture recognition using machine learning is the key to promoting human–computer interfaces that allow users to convey commands and conveniently interact with devices through body postures, hand gestures, etc. In 6G system, the higher-frequency band will enable higher resolution and accuracy to capture finer postures and gestures, and the detection of motion activities (resulting in Doppler shifts) will be more sensitive in the higher-frequency band. Furthermore, the massive antenna arrays allow for recognition with significantly improved spatial resolution and accuracy. Another important benefit of gesture and posture sensing by 6G is the fact that there is no risk of personal privacy information being compromised, as is the case with cameras now, which makes it ideal for many scenarios, especially smart home scenarios. In a future gesture and posture recognition system that utilizes the densely distributed 6G network, devices will be collectively

used to sense the surrounding environments, and sensing data association and fusion at an extended range will significantly improve the overall recognition performance.

There will be advanced gesture and posture recognition features in smart hospitals in the foreseeable future. The medical rehabilitation system in future smart hospitals will enable the automatic supervision of patients. This ensures that their gestures and movements during physiotherapy conform to the standard requirements of rehabilitation exercises. There will be prompt alerts on incorrect movements or gestures, significantly improving patients' rehabilitation. In addition, an alarm alerting the hospital's control center will be generated if a patient falls during an exercise, or if a suspicious person is detected intruding into a restricted area.

The future smart home will be equipped with an advanced hand gesture capturing and recognition system where it allows a hand's 3D position, rotation, and gesture to be tracked. Thus, by simply waving our hands and other gestures, many household appliances such as smart light, smart TV, etc., can be remotely controlled. Looking ahead, more complicated functionalities can be realized by the advanced hand gesture capturing and recognition function in the 6G network, such as playing a virtual piano in the air, in order to provide a completely immersive experience anywhere, anytime. Without doubt, this futuristic concept would open up a range of possibilities for many more innovative applications related to high-accuracy finger motion detection and tracking.

## 2.6 Key Performance Indicators

Within the ISAC context, several new key performance indicators (KPIs) are introduced for sensing capability and they are listed in Table 2.

Table 3 presents the relevant key performance indicators along with the requirements that must be met in order to realize the important use cases discussed in the earlier sections.

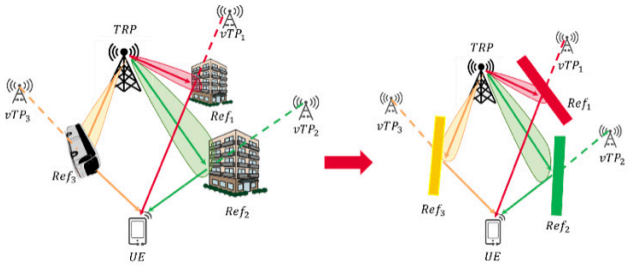
# 3 ISAC for Centimeter-Level Positioning

## 3.1 Background, Motivation and High-Level Scheme

6G requires solutions for sub-centimeter level positioning techniques for various future applications and use cases. This level of accuracy for positioning requires much more detailed knowledge of the radio signal propagation environment where sensing comes into play. By learning the environment RF map and the way the transmitted waveform is manipulated by it, the UE position can be obtained as a function of the measurement parameters. This way, the multipath nature of the propagation channel will be helpful [4]. Moving to higher frequencies can further facilitate such sensing-assisted positioning because the channel becomes sparser, and hence, characterizing the mapping between UE position and its propagation channel takes less effort. In a reflection-dominant environment (which is the case in higher frequencies), one such mapping can be obtained by decomposing the multipath channel as multiple LOS channels coming from multiple anchors. Those anchors are obtained by mirroring the transmission point (TP) over the surface of the corresponding reflector for each path. Those virtual anchors are referred to as virtual TPs or vTPs.

**Table 2** Sensing key performance indicators and their descriptions

| Key Performance Indicator | Description |
|---|---|
| Coverage | Range and field of view limits within which objects can be detected by the system. |
| Accuracy | Difference between the sensed and real values in range, angle, velocity, etc. |
| Resolution | Separation between multiple objects in range, angle, velocity, etc. |
| Detection/False alarm probabilities | Probabilities that an object will be detected when one is present/not present. |
| Availability | Percentage of time for which a system is able to provide the sensing service according to requirements. |
| Refresh rate | Rate at which positioning/localization data is refreshed. |




(a) the real reflectors map (b) the extracted geometrical representation to the visible reflectors

**Figure 1** Mapping the objects/reflectors of the environment to virtual anchors, i.e., mapping multipath components to vTPs

The advantage of such characterization is two-fold:

- The vTPs are totally synchronous with the actual TP. This solves one of the prominent problems of current positioning technologies, which rely on multiple TPs that are not synchronous.
- The channels between vTPs and the UE are LOS, which means that there is no NLOS bias.

Hence, it can solve the two limiting problems of NR positioning (i.e., synchronization error and NLOS error).

However, implementing such technology in a real cellular system is fraught with various challenges and the goal of this section is to provide solutions for these challenges and pave the way for utilizing sensing-assisted positioning in future 6G networks.

- **Potentially large number of vTPs**: In the initial stage of environment sensing, the reflections of the TP location with respect to all objects in the map are obtained. The issue is that number of vTPs grows linearly with the number of reflection planes and grows exponentially

**Table 3** ISAC use cases along with key performance indicators and requirements

| Use Case Category | Coverage | Resolution | Accuracy | Probability | Availability | Refresh Rate |
|---|---|---|---|---|---|---|
| **High-accuracy localization and tracking** | | | | | | |
| Module installation and placement | 10 m | - | 1 mm | - | 99.99% | < 100 ms |
| Docking drone on a moving platform | 50 m | - | 1 cm | - | 99.99% | < 10 ms |
| Robot/Drone as waiter | 50 m | - | 1 cm | - | 99.9% | < 100 ms |
| **Simultaneous imaging, mapping, and localization** | | | | | | |
| SLAM | 50 m | 5 cm | 1 cm | - | 99.9% | < 10 ms |
| Indoor NLOS localization | 100 m | 5 cm | 1 cm | - | 99.9% | < 10 ms |
| Urban environment reconstruction (virtual city) | 100-200 m | 0.5 m | 0.1 m | - | 99% | < 1s |
| **Augmented human sensing** | | | | | | |
| Remote surgery and medical diagnostics | 2 m | 1 mm | < 0.5 mm | - | 99.9999% | < 1 ms |
| Security scans on packages via mobile devices | 0.5 m | 1-2 mm | 0.5 mm | - | 99% | < 100 ms |
| Spectrogram recognition for calories | 0.5 m | 1 mm | 0.5 mm | - | 99% | < 100 ms |
| **Posture and gesture recognition** | | | | | | |
| Medical rehabilitation activity recognition | 10 m | 1 cm | 0.5 cm | - | 99.9% | < 1s |
| Virtual piano anywhere, anytime | 10 m | 0.5 mm | 0.1 cm | - | 99% | < 1 ms |

with the number of allowed bounces. This is, in particular, problematic in outdoor scenarios.

- **Association of the multipath measurements to vTPs**: Another major challenge in implementing the multipath assisted positioning techniques is that a UE has no idea how to match each measurement parameter vector (consisting of angles, delay and Doppler) to a vTP and this can potentially produce a large positioning error. In general, the matching between the observations and the visible vTPs is a combinatorial problem with exponential complexity.

To solve the above issues, we introduce our proposed *sensing-assisted position estimation* (SAPE) scheme. The basic concept of SAPE is to utilize the high resolution capabilities of the massive MIMO and mmWave technologies in space, angular, and time domains in order to increase the resolvability of the multipath components and exploit the environment RF map to identify the potential reflectors of such multipath components, thereby sensing the environment while localizing UEs with high resolution and accuracy. This allows for exploiting the multipath components (including NLOS) to enhance the accuracy of the position, velocity, and orientation information by providing the association between the observations reported from the UEs and the prior information corresponding to the main environment reflectors. Efficient association and accurate mapping need careful design of specific sensing signals, novel transmission and reception signal processing techniques, and their corresponding measurement and signaling mechanisms.

In particular, the proposed SAPE scheme comprises two main steps:

1. **First step sensing or environment sensing**, in which the network (TP) tries to find/update the location of the main reflectors of the environment and obtains the subspace for the next step sensing;

2. **Second step sensing**, in which the TP sends tailored, specific sensing signals in the subspaces obtained in the first step sensing in order to enhance the multipath resolvability and association. The UE performs measurements over the received sensing signals, and by proper mapping of the measurements to the vTPs, the UE position, as well as velocity vector and orientation (which is also referred to as *pose estimation*) can be obtained.

The proposed SAPE scheme is in contrast to most SLAM techniques where all the localization burden/processing is at the UE side.

## 3.2 Detailed Proposed SAPE Scheme

### 3.2.1 First and Second Step Sensing

In the initial environment sensing stage (first step), the TP senses the entire communication space by using a relatively wide beam or small bandwidth in order to generate a coarse RF map to the main reflectors/objects of the communication space. The main goal of this stage is to identify the potential reflectors and map them to vTPs. A static RF map is then available at the TP through this first stage sensing, based on which the location and orientation of the static objects or reflectors can be pre-calculated.

In the second step of sensing, which is the stage of environment sensing update or dedicated sensing, the TP starts targeted sensing based on the obtained RF map and coarse UE location. Particularly, the TP senses certain subspaces, based on the coarse UE location and location of the main reflectors, and processes the reflected signals to obtain finer sensing information of those reflectors. Simultaneously, the UE also performs measurements on the sensing signal to obtain information including multipath identification, range, Doppler, angular and orientation measurements in order to obtain the UE position. Therefore, the second step sensing refines the pre-calculated information obtained in the first phase and thus supports quasi-static environment. In addition, this step can correct the potentially large location errors of the vTP locations obtained from the first step sensing. The impact of vTP location errors will be studied in Section 4.

### 3.2.2 Multipath Parameter Estimation

The problem involves estimating the parameters of the dominant $J$ multipath components of the received signal at the UE per transmitted beam. The parameters to be estimated are the delay $\tau_j$, Doppler $\nu_j$, channel path coefficients $\beta_j$ and angle of arrivals $\vartheta_j^r$, $\phi_j^r$, i.e., elevation and azimuth angles of the $j$-th path. All these parameters are collected into one vector denoted by $\theta_j$, for all $j$. Given the transmitted signal $\boldsymbol{s}_m(t)$ over the $m^{th}$ beam, the received signal is given by:

$$Y_{(m)}(t)=\sum_{j=1}^{J(m)} X_{j,(m)}(t;\theta_j) \quad (1)$$

where $X_j(t;\ \theta_j)$ is the received signal of the $j$-th path. We note here that $X_j(t;\theta_j)$ subsumes the effect of the beamformer at the transmitter and the additive white Gaussian noise at the receiver. We note also that the TX beamforming, during the second step sensing stage, makes $Y_{(m)}(t)$ sparse, i.e., $J_{(m)}$ is small. The joint estimation of these space-time-frequency parameters results in complex noncovex optimization problems. Moreover, the entanglement of the paths' parameters limits the accuracy and reduces the resolution of the estimated parameters, thereby impeding their resolvability. In addition, the high dimensionality in space, time, and frequency, and real-time processing requirements necessitate taking the computational complexity of the parameter estimation algorithm into consideration. Thus, we are looking for a low-complexity super-resolution channel parameters estimator. The literature on the multipath parameters estimation can be classified into four categories, namely, spectra-based [5–6], subspace-based [7–8], compressive sensing-based (sparse signal recovery/reconstruction) and maximum likelihood-based (ML) approaches [9–11]. A high-level comparison between the four categories is provided in Table 4.

Among these algorithms, space alternating generalized expectation (SAGE) maximization is known to be a reasonable approach for reducing the computational complexity, and the slow convergence rate of the maximization step in the EM algorithm is improved by employing the alternating optimization concept over the estimated parameters for each path. Similar to EM, the SAGE consists of two consecutive and iterative steps, i.e., expectation and maximization. In the expectation step, the unobservable data (in our case they are the multipath components $\theta_j$) is estimated based on the observation of the incomplete data and a previous estimate $\hat{\theta}^{(i)}$ of the parameters vector $\theta$. In the maximization step, the parameters vector of $j$-th path $\theta_j$ is re-estimated iteratively by alternatingly optimizing the components of $\theta_j$, i.e., delay, Doppler, channel coefficients and angle of arrivals. In this way, the multi-parameter optimization problem is reduced to multiple single-parameter optimization problems.

### 3.2.3 Multipath Parameter Association

The multipath parameter association problem requires finding a way to associate the estimated parameters, i.e., delays and angles of arrival, of the different channel

**Table 4** Comprehensive comparison between channel parameters estimation

| Category | Description | Pros | Cons |
|---|---|---|---|
| ML-based | Maximum likelihood estimator | The optimal solution | Prohibitive complexity |
| ML-based | Importance sampling ML | Superior performance | High complexity<br>Slow convergence |
| ML-based | Expectation maximization | Superior performance | High complexity<br>Slow convergence |
| ML-based | SAGE | Super resolution | Medium complexity<br>Fast convergence |
| Sparse signal reconstruction | OMP and its variant | Competitive performance | Medium complexity |
| Sparse signal reconstruction | Based on convex relaxations such as $l_1$ norm, nuclear norm, and atomic norm | Super resolution | High complexity |
| Subspace-based | MUSIC, ESPRIT, and their variants | Medium resolution | Medium complexity |
| Subspace-based | FFT | Low resolution | Low complexity |

multipath components, $Z = \{\hat{\theta}_1,..,\hat{\theta}_{N_{obs}}\}$, where $N_{obs}$ denotes the number of observation and $Z$ is obtained by the UE for each measurement-beam pair, to their relevant visible vTPs represented by the set of the ground truth values $G = \{\boldsymbol{g}_1,..,\boldsymbol{g}_{N_{vtp}}\}$, where $N_{vtp}$ denotes the number of visible vTPs and $G$ is obtained by the TP through the first and second step sensing, as shown in Figure 2.

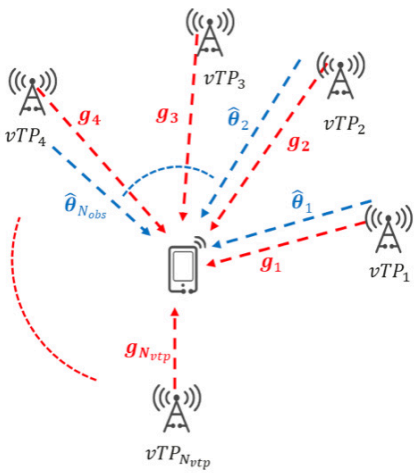



**Figure 2** Illustration of measurements — vTP association problem

Extensive research has been conducted in order to alleviate the association error and reduce the computational complexity of the association algorithms. These works can be categorized into two main lines of thoughts, namely, soft-decision/probabilistic data association and hard decision data association [12]. In the probabilistic approach [13–16], all the vTPs are assigned to a certain measurement/observation with different probabilities, with the probabilities indicating how likely a given measurement is due to a particular vTP. This requires a proper selection of the statistical model for assigning these probabilities. In the hard decision data association approach [12, 17], each measurement/observation is associated only to one vTP. The techniques within this approach can be divided into two categories, namely, probabilistic-based hard decision algorithms and distance metric-based selection algorithms. In the former, the measurement is associated to the most likely association event according to a certain probabilistic measure such as maximum likelihood or a posteriori; in the latter, the measurement is associated to the nearest association event according to a certain distance metric such as the Mahalanobis distance [12, 17]. The main drawback of this approach is that it depends heavily on the accurate knowledge of the UE position as a prior.

Prior art adopts measuring the distance between the set of the measurements $Z = \{\hat{\theta}_1,.., \hat{\theta}_{N_{obs}}\}$ and the set of the expected ground truth values $G = \{\boldsymbol{g}_1,..,\boldsymbol{g}_{N_{vtp}}\}$, where $\hat{\theta}_i$ is the $i$th vector that contains range and angles of arrival of the measurement and $\boldsymbol{g}_k$ is the distance vector between the expected UE position $\hat{\boldsymbol{p}}$ and the $k$-th vTP obtained from previous estimations. However, this technique has two main shortcomings. First, it requires prior information about the UE's position which might not be available in many scenarios. Second, it requires using long training periods and measurements to iteratively update the prior knowledge of the UE's position to make the association algorithm converge. To cover these shortcomings, we propose a new technique that mainly exploits the differential/mutual distances between the members of the two measurement sets. The key idea of the proposed algorithm is to match the relative/differential distances between the members of the measurements set to the relative distances between a subset of visible/expected vTPs as shown in Figure 3.

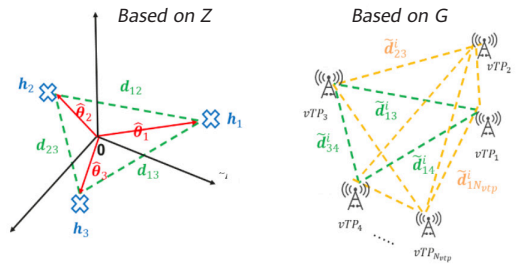



**Figure 3** Illustration of relative/tdifferential distance (a) based on $Z$ (b) based on $G$

This is mainly aimed at avoiding the dependency of $G$ on $\hat{\boldsymbol{p}}$. This requires two different modifications on the two sets, $Z$ and $G$. First, instead of directly using the $N_{obs}$ measurements, we convert them into $N_{obs}$ hypothetical vTP locations relative to the origin point. We denote this set of hypothetical vTP locations by $H_z = \{\boldsymbol{h}_1,..,\boldsymbol{h}_{N_{obs}}\}$, where $h_i$ is the relative location vector of the $i$-th hypothetical vTP. Based on these relative location vectors, we calculate the differential/mutual distances between these relative locations, i.e., $\boldsymbol{d}_{ij} = \boldsymbol{h}_i - \boldsymbol{h}_j \forall_i, j, i \neq j$. The set of the differential/mutual Euclidean distances between the hypothetical vTPs' locations is defined as $D$ and has a size of $\binom{N_{obs}}{2}$ elements where its $n$-th element is denoted by $\boldsymbol{d}_{ij}(n)$. The second modification is to build the set of the differential/mutual distances between the real vTPs' locations, i.e., $\tilde{D}$. Because the set of hypothetical vTPs' locations, i.e., $H_z$ and the set of the actual vTPs locations, i.e., $R$ usually have a different cardinality, we divide the later set into $\binom{N_{vtp}}{N_{obs}}$ subsets of size

$N_{obs}$, with each containing a different combination of vTPs' locations. We denote these subsets by $r_i = \{\boldsymbol{r}_{i1,\ldots,}\boldsymbol{r}_{iN_{obs}}\}$, $i \in \{1,\ldots,\binom{N_{vtp}}{N_{obs}}\}$, where $\boldsymbol{r}_{i1}$ is the location vector of first vTP in the $i$th subset. For each $r_i$, we define the differential/mutual distances between its members as $\tilde{\boldsymbol{d}}^i_{mn} = \boldsymbol{r}_{im} - \boldsymbol{r}_{in}, \forall m,n,m\neq n$. The set of the mutual/differential between the members of $r_i$ is denoted by $\tilde{D}^i$. We measure the distance between the sets $D$ and $\tilde{D}^i$ by:

$$d^i_{DD} = \min_{\pi_j} \sum_{n=1}^{\binom{N_{obs}}{2}} \left| d(n) - \tilde{d}^i_{D_{\pi_i}}(n) \right| \quad (2)$$

where, with a slight abuse of notation, $d(n)$ and $\tilde{d}^i_{\tilde{D}^i_{\pi_j}}(n)$ are $n$-th elements of $D$ and $\tilde{D}^i_{\pi_j}$, respectively, and $\pi_j$ is the $j$-th permutation of the elements of $\tilde{D}^i$. Using this new metric, the association is given by:

$$\pi_{opt} = \arg_i \min d^i_{DD} \quad (3)$$

$\pi_{opt}$ contains the indices of those entries of $R$ that are optimally assigned to $D$.

## 3.3 Performance Evaluation of the Proposed SAPE Scheme

### 3.3.1 Link-Level Evaluation

In this subsection, we mainly evaluate the average behaviour of the proposed association algorithm using statistically generated measurements. In the simulation, we generate 8 uniformly distributed vTP positions. We further assume each received observation $\theta$ (can be range or angle) has Gaussian noise with standard deviations of $\sqrt{c*\mathrm{CRLB}(\theta)}$, where CRLB($\theta$) denotes the Cramer-Rao Lower Bound for mean square error estimate of $\theta$ and the constant factor $c$ accounts for the non-ideal factors in the detection which results in the gap between the real estimation and the lower-bound (in the evaluation results, $c$ = 6). We calculate the association error by counting the number of different indices of the associated vTPs from the observed ones.

As shown in Figure 4, the proposed association algorithm provides a very good performance in the synthetic scenario. In addition, better performance is observed with more measurements (more vTPs) because more mutual/differential distances are used for association. We note that the association error in Figure 4 represents the ratio of the number of vTPs wrongly associated to the total number of vTPs on average. For instance, in Figure 4, it is shown that with 3 observations, one gets an average association error of 0.07 at SNR of 5 dB, i.e., just 7 out of 100 vTPs on average will be associated wrongly. It is also noteworthy to mention that the average association error is different from the average position error. However, the former affects the later. In other words, wrongly associating one out of 10 vTPs might not produce significant position error if the measurements for this vTP have a lower weight in calculating the position error.

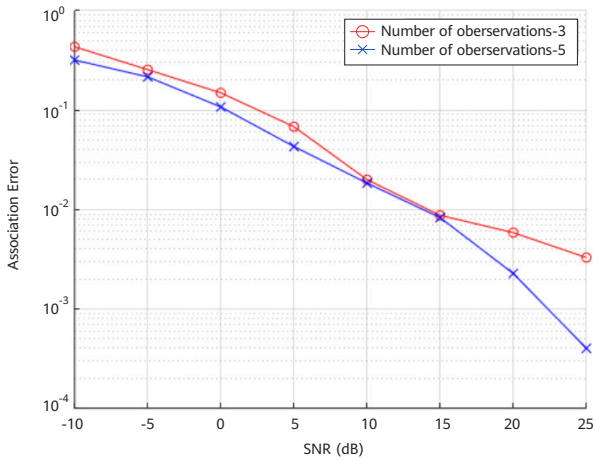



**Figure 4** Average association error of the proposed association algorithm when the number of visible vTPs is 3 and 5

We further evaluate the performance of the proposed association algorithm and the impact on the positioning error in a real multipath environment. Without generality, we assume that the estimation error due to the channel parameters estimation stage follows the CRLB. Figure 5 presents the CDF of the UE positioning error bound (PEB) due to the association scheme and compares it with the case of ideal association, which shows the promising performance of the proposed SAPE technology and its potential for 6G positioning.

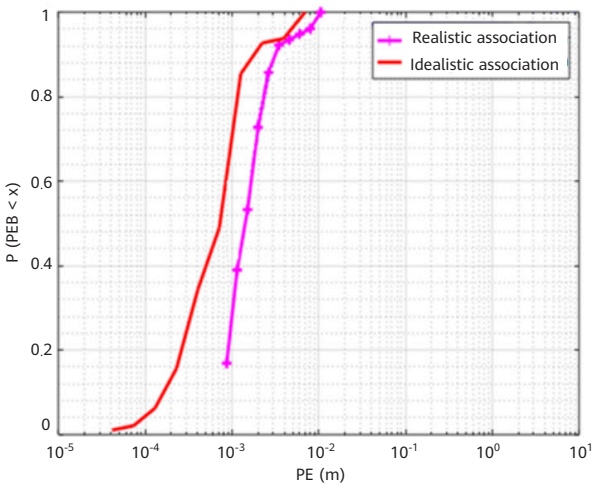



**Figure 5** Overall positioning performance of the proposed SAPE technology at the link level

### 3.3.2 System-Level Evaluation

In this section, we provide the performance of the proposed SAPE scheme at the system level. Similar to the system-level communication performance evaluation, the key step in such an evaluation is abstracting the PHY-level performance at the network level, which is the so-called PHY abstraction. The rationale behind sensing PHY abstraction is to map the system parameters in terms of SINR, bandwidth, time duration, and antenna configuration to a sensing performance (i.e., range, Doppler or angle mean square errors). The proposed SAPE scheme is evaluated and compared with baseline NR in terms of PEB, based on the proposed PHY abstraction methodology in two scenarios:

*Idealistic scenario*: where the sensing is assumed to be perfect. In this case, the evaluation is based on applying the proposed PHY abstraction methodology in SLS and evaluating the candidate schemes in two scenarios: indoor hotspot (InH) and outdoor urban micro (UMI). Both scenarios are evaluated over mmWave bands and the simulation parameters are given in Table 5.

**Table 5** Parameters for SLS evaluation

| Parameter | Value |
|---|---|
| Bandwidth | 80 MHz |
| Sensing time | 14 symbols |
| Sub-carrier spacing | 60 kHz |
| Number of subcarriers | 1024 |
| Deployment | Indoor hotspot, 256 × 32<br>UMI 32 × 16 (outdoor only),<br>20 RRUs and 200 UEs |
| Channel model | SCM (stochastic) |
| Carrier frequency | 60 GHz |
| Simulation methodology | Based on SLS using the proposed sensing PHY abstraction |
| Non-idealities modeled | Sensing error |
| Synch. error between TRPs | 0 (perfect synch.) or 1 ns |

Based on these parameters, the simulation results are given in Figure 6.

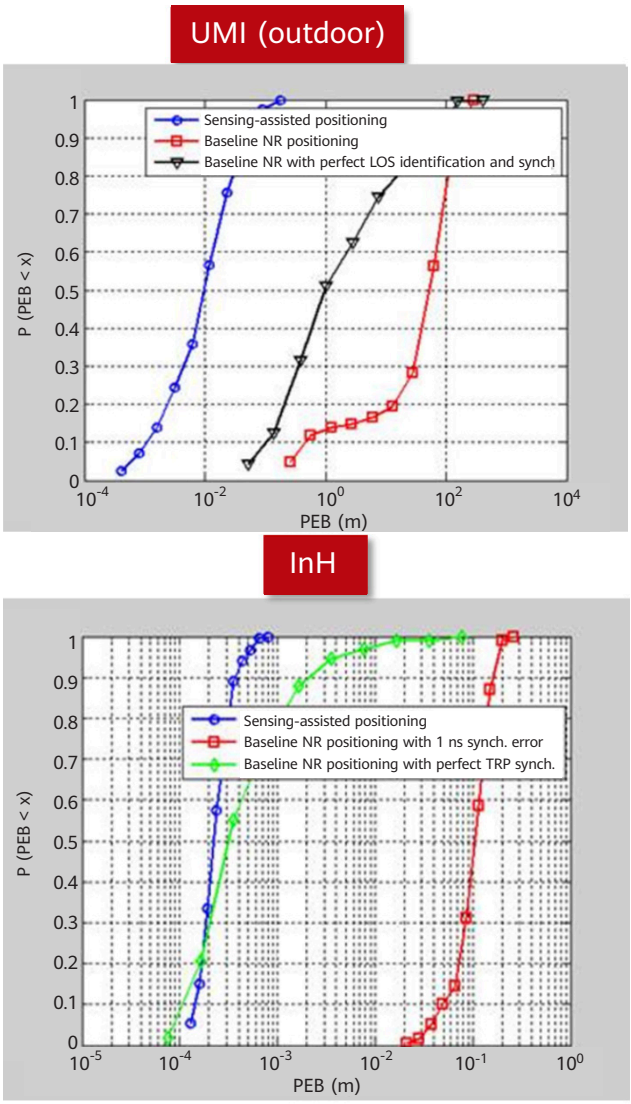



**Figure 6** SLS results of the proposed SAPE vs. baseline NR in idealistic scenario

Based on the results, we can observe that under ideal conditions (no RF impairments, no sensing error, no diffraction), SAPE can achieve an order of magnitude better accuracy compared to NR. In addition, the NR baseline cannot achieve good performance in any scenario, even under ideal conditions, due to NLOS bias and synchronization error between the TPs.

*Realistic scenario*: assuming sensing error, the candidate schemes are evaluated in outdoor UMI. The simulation parameters are the ones given in Table 5. For modeling the sensing error, we assume the vTPs corresponding to each path/cluster are Gaussian-distributed with some variances which are also modeled as random variables. In addition, the vTP location variance for the LOS link is set to 0 as it corresponds to the actual TP. Based on these parameters, the simulation results are given in Figure 7.

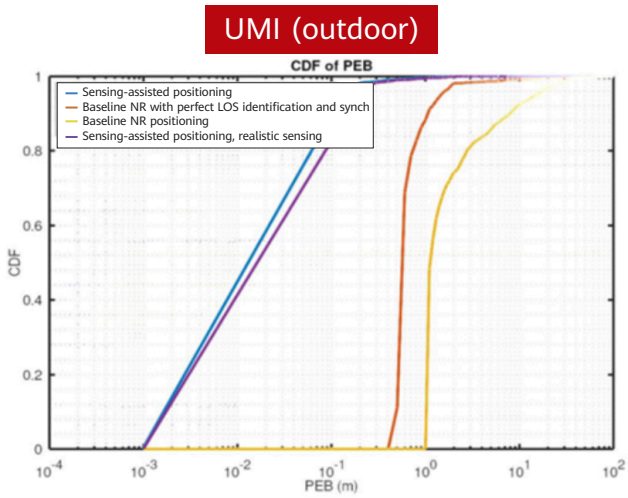



**Figure 7** SLS results of the proposed SAPE vs. baseline NR in the realistic scenario

Based on the results, we can observe that SAPE can achieve an order of magnitude better accuracy (cm-level accuracy) when compared with NR, even with the sensing error.

## 4 ISAC for Millimeter-Level Imaging at the THz Band

THz lies between the mmWave and infrared frequencies, and thus has millimeter-level and even sub-millimeter-level wavelength, making the ISAC system at the THz band (ISAC-THz) particularly suitable for high resolution sensing applications such as millimeter-level resolution 3D imaging. Like the other lower frequency radio waves, THz can penetrate some obstacles, achieving high-precision sensing in all weather and lighting conditions.

Recent developments in semiconductor technology have bridged the "THz band gap" and made the hardware feasible at the terminal side. ISAC-THz based portable devices will thus open the door for numerous new sensing applications such as augmented human sensing with very high resolution. Table 6 shows the allocated mobile frequency bands with a contiguous bandwidth greater than 5 GHz. The ultra-wide bandwidth in THz will also enable Terabits/second data rate transmission, especially in short-range communications. The corresponding range resolution (from equation 4) based on Heisenberg's Uncertainty Principle is also provided in this table. Under the assumption of synthesized aperture, the cross-range resolution is provided in Table 7 based on equation 5 where $\lambda$ is the wavelength, $D$ is the aperture size, and $r$ is the distance between transceiver and target. Because THz can provide high sensing resolution in addition to high communication throughput, the integration of THz sensing and communication has become an attractive and active research area.

$$\triangle t = \frac{C}{2B} \tag{4}$$

$$\triangle d = \frac{\lambda r}{2D} \tag{5}$$

The application of ISAC-THz design is expected to provide many opportunities for brand new services especially on future mobile devices or even wearables as illustrated in Figure 8. In addition to the localization and imaging applications, molecular spectrogram analysis is another interesting application area that could be enabled by ISAC-THz, as discussed in Section 2.

**Table 6** Maximum contiguous bandwidth in the range of 100-450 GHz and the corresponding range resolution

| Freq. (GHz) | Contiguous Bandwidth (GHz) | Range Resolution (mm) |
|---|---|---|
| 102–109.5 | 7.5 | 20 |
| 141–148.5 | 7.5 | 20 |
| 151.5–164 | 12.5 | 12 |
| 167–174.8 | 7.8 | 19 |
| 191.8–200 | 8.2 | 18 |
| 209–226 | 17 | 8.8 |
| 252–275 | 23 | 6.5 |
| 275–296 | 21 | 7.1 |
| 306–313 | 7 | 21 |
| 318–333 | 15 | 10 |
| 356–450 | 94 | 1.6 |

**Table 7** Aperture size and corresponding cross-range resolution at 140 GHz

| λ = 2.1 mm, r = 30 cm | |
|---|---|
| **Aperture Size (cm)** | **Cross-Range Resolution (mm)** |
| 1 | 32 |
| 5 | 6.4 |
| 10 | 3.2 |
| 20 | 1.6 |

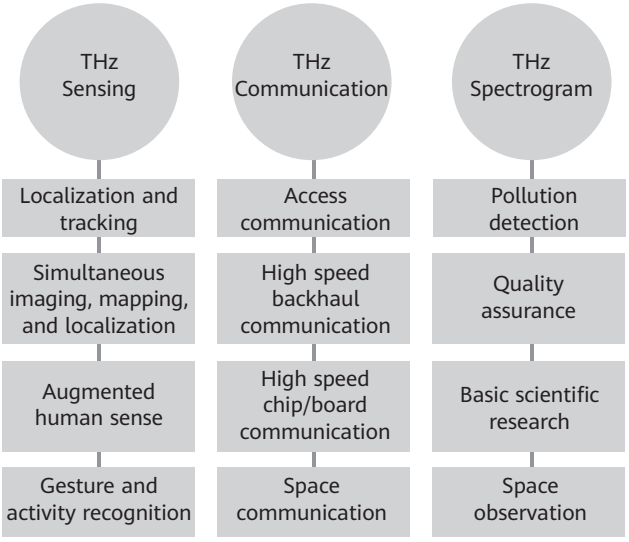



**Figure 8** THz application in sensing and communication

In this section, we elaborate on our ISAC prototype of THz imaging on portable devices that achieves millimeter-level resolution. A robot arm equipped with ISAC-THz module is used to represent a human arm holding a THz imaging camera. The prototype is built to operate at 140 GHz carrier frequency with a bandwidth of 8 GHz.

## 4.1 Hardware Architecture of the ISAC-THz Module

From the THz imaging aspect, thousands of antenna elements are required to create a large aperture for high cross-range resolution. However, it is clear that physically packing thousands of antenna elements into the portable device is infeasible due to the size and power constraint requirements of the device [18–19]. To solve this problem, virtual aperture techniques are applied in the prototype system [3]. In particular, the virtual MIMO antenna array design in the hardware transceiver architecture using the sparse sampling design in the scanning process is proposed [3, 20–21].

First, a virtual MIMO antenna array structure is constructed to form a virtual aperture that can achieve the same performance with respect to its equivalent physical aperture array as illustrated in Figure 9. Next, a sparse scanning approach is applied, transforming the degree-of-freedom in time and space into a larger virtual aperture, as shown in Figure 9. The scanning performed by the robot arm thus mimics a user holding a smartphone and imaging an object with a zigzag scanning trajectory.

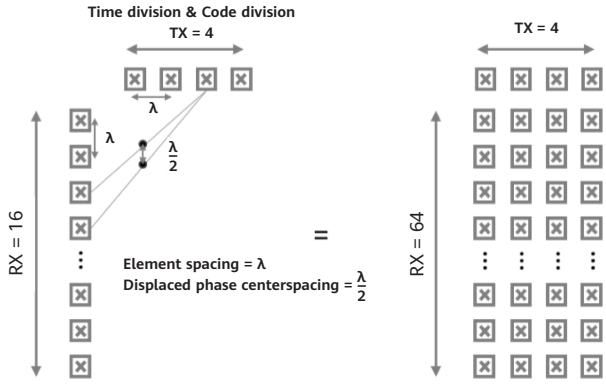
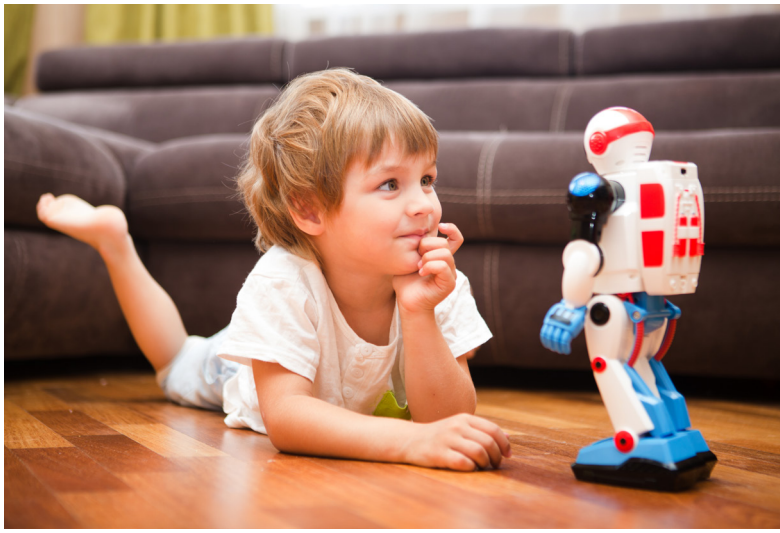



**Figure 9** MIMO virtual aperture

To implement the overall solution of a virtual aperture, the following three requirements need to be satisfied in the hardware design:

- Multiple transceiver (TRX) chains to support the MIMO antenna array structure as the first step for the overall virtual aperture.
- Wide antenna pattern to cover the target scanning area in order to maintain the correlation among the reflected samplings.
- Real-time position information of the device to perform coherent processing of the received signals.

The schematic of the prototype architecture is shown in Figure 10. The transmitter antenna array has 4 RF ports and the receiver antenna array has 16 RF ports, forming a 4T16R MIMO antenna array structure [3]. The per unit antenna radiation pattern is a wide beam design with a 3 dB beam width of 50° and gain of 7 dBi.



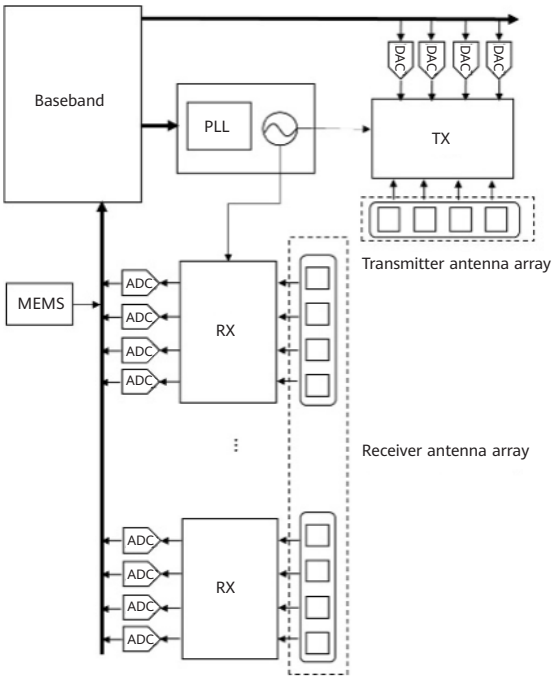



**Figure 10** Illustration of the architecture of ISAC prototype

## 4.2 Compressed Sensing-based Tomography Imaging

A major challenge for the virtual aperture imaging technique is the irregular scanning trajectory caused by the user moving the ISAC imaging module to perform THz scanning on an object. Assume a zigzag scanning routine is used to image an object, as shown in Figure 11. The echo samplings in the horizontal direction are continuous, i.e., the spatial spacing between sampling points is comparable to the wavelength of the echo signal. However, continuous sampling cannot be maintained in the vertical direction. As a result, the echo samplings in the vertical direction are sparse, which will cause high and non-uniform sidelobe effects, giving rise to false artifacts, which may lead to imaging failure.

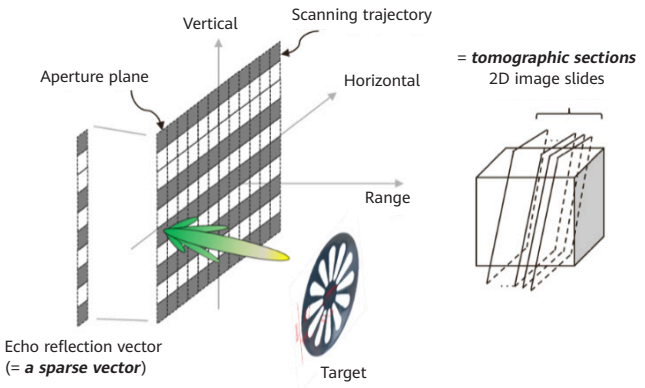
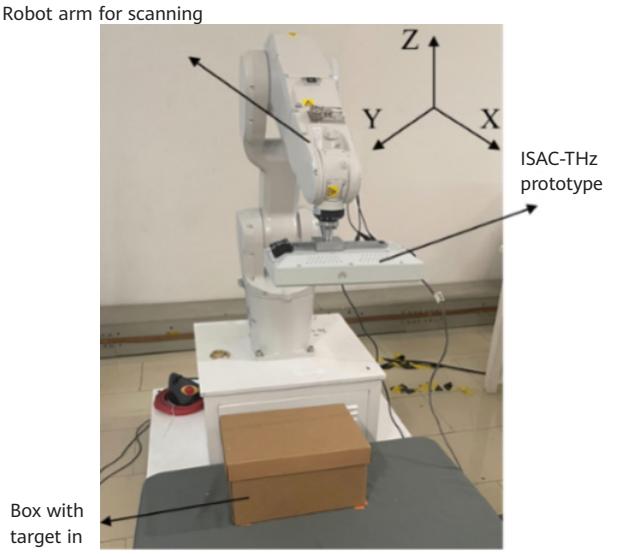


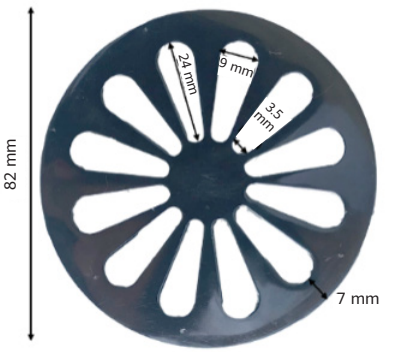

**Figure 11** Illustration of the sparse scanning approach and the tomographic imaging techniques

To solve this challenge, we consider decomposing the scanning trajectory on a two-dimensional (2D) plane into several sets of linear scanning tracks along the horizontal direction, where the sparseness of the sampling signals in the vertical domain is then equivalent to the sparseness between horizontal tracks, as illustrated in Figure 11. In this case, the reflected/echo information from the object can be retrieved from these vertically sparse samplings via compressed sensing techniques [3].

As depicted in Figure 12a, the robotic arm scans at a speed of 1 m/s with the scanning area set as 10 cm by 12 cm in the prototype. The longitudinal spacings of the scan trajectories are controlled to simulate the sparsity in the trajectories of the user's hand-held scanning behavior. The target object to be imaged, as shown in Figure 12b, is put in a box with a cap on top of it. As we can see from Figure 12b, the smallest distance in the hallowed pattern is 3.5 mm, so the highest resolution of the imaging results can be 3.5 mm.




(a) Prototype setup for THz sensing where the ISAC-THz module is held by a robot arm representing a human arm

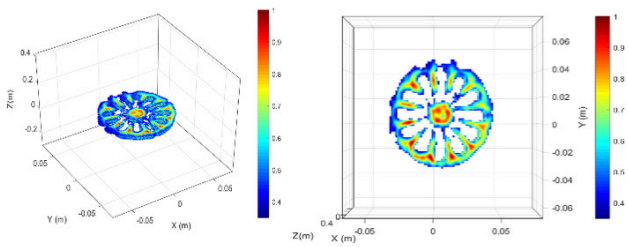



(b) Target object in the box

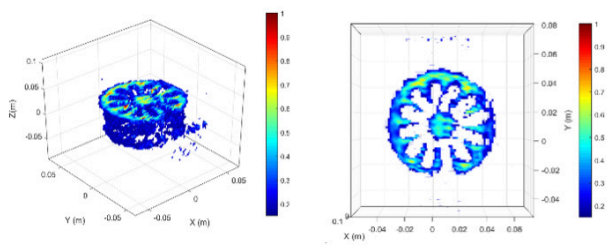
**Figure 12** Setup of the ISAC-THz prototype

The proof-of-concept THz imaging performances with different sparsity configurations in the scanning patterns are presented and compared in Figure 13. In each of the figures, the 3D imaging results are shown on the left and the cross-range profile perceived from top down is shown on the right.

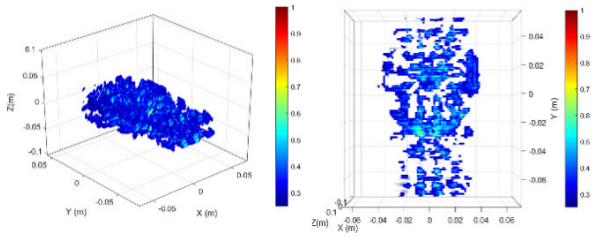
The non-sparse full aperture scanning in Figure 13a is an ideal case, in which the vertical sampling is half wavelength adjacent. This achieves the best PSLR and ISLR performance, which is set as an upper bound performance reference. Then, in order to simulate the sparsity in real free hand scanning, we assume different sparsity configurations in tests, from 50% (medium sparsity) to 25% (most sparsity), where *X*% sparsity means that there are *X*% of the full samplings remaining in the vertical direction. With the collection of fewer samplings, stronger side-lobe interference occurs at the resulted aperture, resulting in worse imaging performance. From the comparison of Figure 13c and Figure 13d, we see that when the sparsity is too high, the traditional tomography algorithm is not enough to recover the images. In this case, the compressed sensing based tomography approach showed its superior performance.



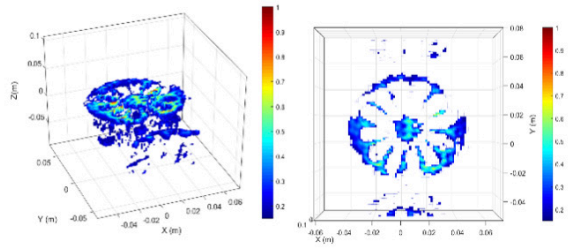
(a) Non-sparse full aperture scanning (ideal case)



(b) Sparse scanning with 50% sparsity (medium sparsity)



(c) Sparse scanning with 25% sparsity (most sparsity) and using the traditional tomography approach [22]



(d) Sparse scanning with 25% sparsity (most sparsity) and using the compressed sensing based tomography approach

**Figure 13** Imaging results at different sparsity configurations

## 4.3 Multi-Channel Imaging

The multi-channel imaging process can be treated as a time-domain coherent combination of electromagnetic signals from multiple receiving channels. Theoretically, n receivers can reduce the sampling time to 1/n compared with one receiver with the same imaging quality. Less sampling time will reduce the difficulty of motion error compensation, which in turn will improve the imaging quality.

However, in multi-channel imaging, one major challenge arises from the imbalance in gain and time delay of different receiver channels due to hardware imperfection. The antenna mounting positional imperfection will introduce the displaced phase center error, as shown in Figure 14. Multi-channel amplitude and phase imbalance will lead to azimuth ghosting, which will significantly degrade the imaging quality. The amplitude imbalance can be easily compensated by multichannel amplitude equalization methods [23], while phase imbalance compensation needs auto-focusing algorithm such as gradient descent.

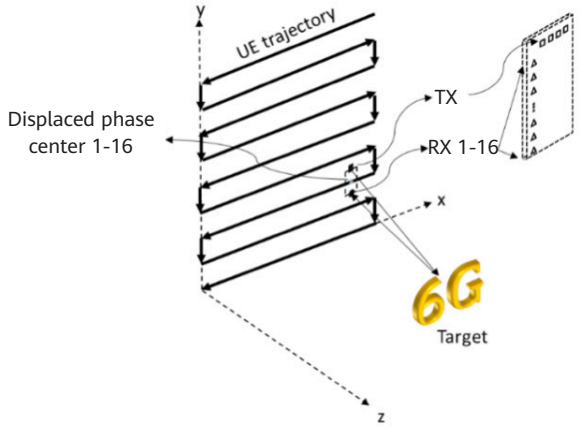
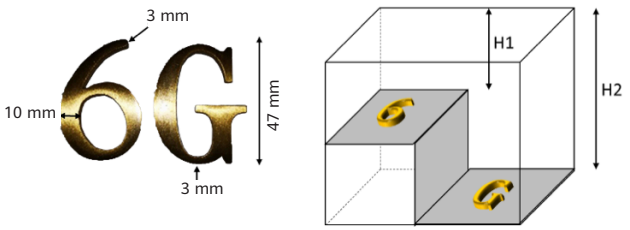



**Figure 14** Illustration of the displaced phase center caused by multi-channel imaging

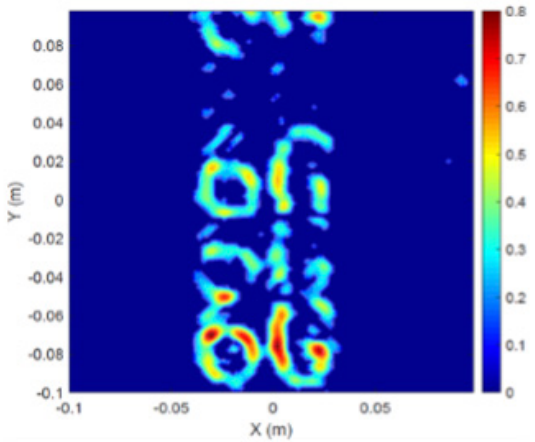
To validate the performance of multi-channel imaging, we use the same testbed described in the last subsection but a different target (resolution is similar, i.e., 3 mm) as shown in Figure 15. In this case, we tried both the 2D target shown in Figure 15a and the 3D target shown in Figure 15b, where the 3D imaging target was formed by placing the two characters at different heights inside the box. With the aforementioned benefit of multi-channel imaging, very sparse sampling is needed for a good imaging quality. In the prototype, only 12% sparsity is configured in the scanning trajectory.



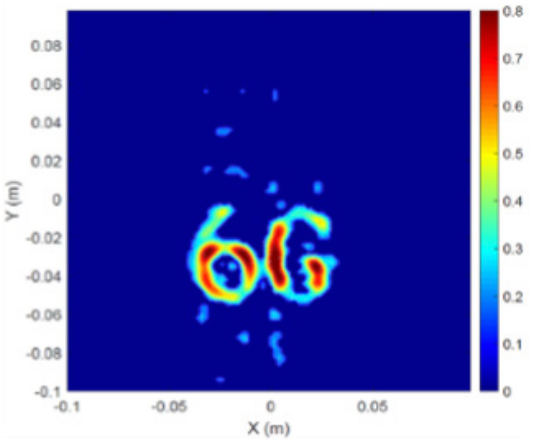

(a) 2D imaging target (b) 3D imaging target

**Figure 15** Imaging targets used in multi-channel imaging

Figure 16 shows the imaging results of the 2D target. The imaging results without multi-channel phase error compensation are illustrated in Figure 16a where severe ghosting on the final image that significantly degrades the imaging quality can be prominently seen. Using the geometric interpretation algorithm, the sidelobe due to the multi-channel imbalance has been duly suppressed as can be clearly seen in Figure 16b.




(a)




(b)

**Figure 16** Multi-channel imaging result of the 2D target

Subsequently, the imaging results of the 3D target are shown in Figure 17. The imaging result clearly depicts the shape of the two characters and their relative distance in the 3D space.

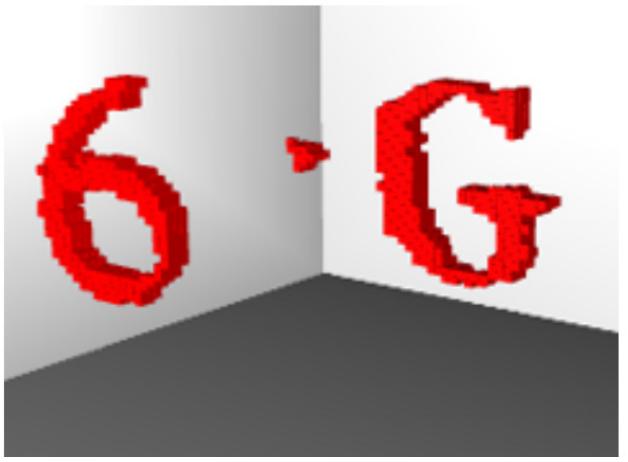


**Figure 17** 3D imaging result

# 5 Major Challenges for Making ISAC a Reality

## 5.1 Channel Modeling and Evaluation Methodology

In 6G, the channel model needs to be considered for both communication and sensing services. This brings significant challenges to the channel modeling methodology. Until 5G, because it has low computational complexity and is easily standardizable, stochastic channel modeling methodology dominated the evaluation of wireless communications, and is used in many projects and standards such as 3GPP-SCM, WINNER-I/II, COST2100, and MESTIS. It is adequate in evaluating the communication performance. However, there is a doubt as to whether it still can meet the more diverse requirements from different sensing applications.

One typical sensing channel is the echo channel, which consists of the backscattering RCS characteristics from the object and its surroundings. This type of propagation channel brings new requirements for the physical electromagnetic (EM) characteristic which are not supported in the current communication channel models. One typical use case is the high resolution imaging application. This type of application requires the deterministic channel coherence of the antenna array aperture with the geometry information. This requirement is contradictory to the typical stochastic channel modeling approach. Therefore, the traditional channel modeling methodologies deserve some rethinking and innovation.

Another major challenge is the evaluation performance metrics based on the new sensing requirements. Conventionally, throughput, latency, and reliability are the main evaluation performance metrics for communication systems. However, due to the different sensing applications, there are new dimensions of evaluation metrics that need to be considered, such as sensing resolution, accuracy, detection probability, and update rate. So far, no KPIs have been proposed for the joint performance characterization and evaluation of both the communication and sensing services. This implies that a new scenario-dependent evaluation methodology may need to be investigated.

To address the challenges mentioned above, the following research directions are proposed:

- Typical scenarios and evaluation methodology

The traditional indoor hotspot, urban micro, urban macro are defined in the 3GPP 38.901 communication channel. The environment and the purpose of the application will deeply affect the channel model parameters and even the channel model generation approach. Therefore, the ISAC typical scenario should be categorized and the typical use cases of each category should be highlighted for further evaluation.

For a given evaluation use case, metrics to characterize the joint performance between communication and sensing are needed in order to optimize the performance trade-off for both services simultaneously. To characterize the performance of both functions as well as mutual enhancement, scenarios and metrics need to be implemented into the system level simulations and the ISAC performance must be evaluated in a fully integrated network.

- Channel measurement and modeling methodology

Regarding ISAC channel modeling, a single channel modeling scheme may not meet the need to evaluate all ISAC applications. Instead, stochastic, deterministic, and even hybrid channel models must be considered. For instance, in the sensing-assisted beamforming use case, the stochastic channel modeling could be adopted, whereas for localization and tracking application, ray tracing could be considered as a strong candidate for channel modeling because the detailed contours for the object reflection/scattering are not strictly required. On the other hand, for imaging and recognition applications, there is a need to consider EM algorithm when the size of the scatterers is close to the signal wavelength and therefore the interaction of the signal to the scatterers are strongly correlated with the EM characteristics.

## 5.2 Joint Waveform and Signal Processing Design

Most of the works on the joint design of sensing and communications mainly focus on the joint waveform design. The main challenge for the joint waveform design is the contradicting KPIs for communications and sensing. In particular, the main target for communications is maximizing the spectral efficiency, whereas the optimum

waveform design for sensing is focused on estimation resolution and accuracy. Because CP-OFDM has been proven to be a favorable option for communication, many researchers have considered this waveform for sensing as well. Although the introduction of cyclic prefix (CP) has been shown to degrade auto-correlation in the time domain [24], a novel approach of frequency domain processing [25] allows for efficient parameter estimation of CP-OFDM, achieving the maximum processing gain. Furthermore, CP-OFDM has been shown to be free of the range-Doppler coupling problem, which means that the range and Doppler estimation can be performed independently [25]. However, these favorable properties of CP-OFDM depend on perfect synchronization (in both time and frequency domains) between the transmitter and the receiver, which may not be present, especially for bistatic sensing. In addition, the large peak to average power ratio (PAPR) of CP-OFDM is another major issue for radar applications where power efficiency is very important.

Alternatively, frequency modulated continuous wave (FMCW) waveform, which has traditionally been used for radar, is not capable of carrying data at transmission rates desirable for communication services. Some researchers proposed to modify the FMCW waveform to make it more communication-friendly. Among the many contributions in this line of research, we can mention [26], in which the authors propose to use up-chirp for communication and down-chirp for radar, and [27], introducing trapezoidal frequency modulation continuous-wave (TFMCW) modulation in which the radar cycle and communication cycles are multiplexed in the time domain. Although these techniques enable efficient multiplexing of communication data in the sensing signal, they still suffer from low spectral efficiency due to the existence of the chirp-like sensing signal. Another line of research is devoted to using single-carrier waveform based on the code domain spreading of joint radar and communication signals. For this class of waveforms, the radar performance has been shown to be affected by the auto-correlation of the sequences and the long spreading codes result in good auto-correlation at the expense of communication spectral efficiency [27]. In addition, Doppler estimation requires more complicated algorithms [27]. The current state of the art suggests that there is still room for waveform design to strike a balance between good communication and sensing performance to meet 6G ISAC requirements.

## 5.3 Hardware Co-design

In the design of the ISAC system, the solution that integrates the baseband and RF hardware reduces the overall power consumption, system size, and information exchange latency between the two systems. The hardware converging strategy facilitates the mutually beneficial functions of sensing and communication in distortion calibration and compensation. The common impairments in the ISAC system due to hardware imperfections are demonstrated in Figure 18.

It should be noted that in light of the differences in evaluation metrics and algorithms between communication and sensing, hardware requirements are quite different. Considering the cost and size of the historical communication and radar systems, the ISAC system hardware design will closely resemble the traditional communication architecture. As a tradeoff, we need to consider the impact of distortion parameters on sensing performance. For instance, a communications system depends on full duplex isolation to achieve high capacity.

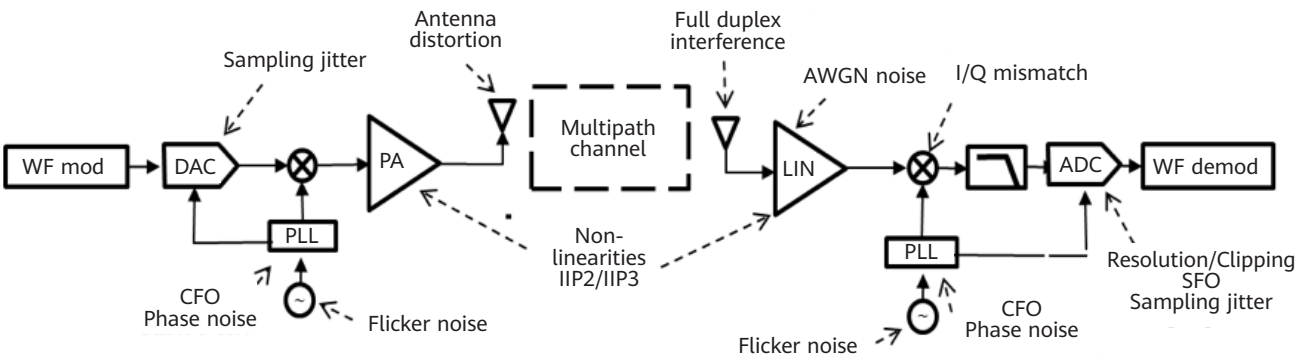



**Figure 18** Impairments of ISAC transmission system

In contrast, from the OFDM ISAC perspective, limited transmitter-receiver isolation is a primary concern in the detection of static targets [28]. Proper design of integrated RF architecture and self-interference cancellation in the ISAC system are key technical problems that need to be solved. Another issue is that sensing requires the accumulation of coherent signals to ensure performance, which makes the system more sensitive to sampling jitter, frequency offset, and phase noise [29]. This in turn leads to higher requirements on synchronization and stability of the system. In short, we need to consider these hardware challenges in the selection of ISAC waveforms, sensing algorithms, and non-ideal distortion compensation schemes.

## 5.4 Sensing-assisted Communication

Although sensing will be introduced as a separate service in the future, it might still be beneficial to look at how the information obtained through sensing can be used in communication. The most trivial benefit of sensing will be environment characterization, which enables sensing-assisted communication due to more deterministic and predictable propagation channels. It has been shown that the environment knowledge provided by sensing not only improves the accuracy of channel estimation in mmWave, but also significantly reduces the overhead because the environment is shared by potentially many UEs and sensing-based channel acquisition does not repeat the channel estimation process for each individual link [30]. Another example would be sensing-assisted beam alignment, especially in mmWave vehicular communication where the main challenge is the frequent link reconfiguration resulting in significant overhead. In [31], it has been proposed to use the information obtained from a radar mounted on an infrastructure operating in a given mmWave band to configure the beams of the vehicular communication system operating in another mmWave band. Moreover, the users' location information and the environment map obtained by sensing helps identify the link blockage caused by large objects, especially in dense urban networks, so that the power and beams can be adjusted accordingly to improve the communication throughput [32]. Other examples of sensing-assisted communication can also be considered and studied to reduce the latency and overhead of communication systems in future 6G networks with the help of information provided by the sensing system. Another benefit of sensing for communication would be improving the users' positioning accuracy by combining the advantages of active localization and passive localization and thus overcoming their shortcomings to satisfy 6G localization requirements.

## 5.5 Communication-assisted Collaborative Sensing

6G ISAC takes advantage of the mobile communication network to support synchronized, collaborative multi-node sensing. Sensing through cooperation refers to the sensing nodes that share their observations with each other and attempt to reach a common consensus on the surrounding environment. This will significantly improve localization performance. Integration of sensing capabilities into the existing communication network will be the most viable and cost effective option where the multiple network nodes (base stations, UEs, etc.) can function as a complete sensing system to enable network sensing operations for the use cases highlighted in Section 2. The process involves the collaborating nodes forming a dynamic reference grid through distributed sensing and processing. The collaboration reduces measurement uncertainty and provides greater coverage as well as higher sensing accuracy and resolution through sensing data fusion. In addition, this offers interesting possibilities for being able to carry out sensing under non-line-of-sight (NLOS) conditions. The major research challenges here would lie in the synchronization, joint processing, and network resource allocation in order to achieve the optimum sensing fusion results.

# 6 Conclusion

With the concept of ISAC being commonly accepted as one of the key technology trends for 6G, this paper takes a step forward and elaborates two case studies on how 6G ISAC technologies can be applied to improve localization and to perform high resolution imaging. In particular, the proposed SAPE scheme utilizes the joint benefit of device-free and device-based sensing and greatly improves the positioning accuracy compared with the current NR scheme. The prototype of the THz camera justifies the feasibility of mm-level imaging resolution on portable devices for

both 2D and 3D objects placed in a box. Joint efforts from both academia and industry are needed to address further challenges in the system level evaluation of ISAC, new channel modeling methodology, new waveform design, low complexity algorithm design, and low cost hardware design.

## References


[1] W. Tong, P. Zhu, *et al*., "6G: the next horizon: from connected people and things to connected intelligence," Cambridge: Cambridge University Press, 2021.

[2] D. K. P. Tan, J. He, Y. Li, A. Bayesteh, Y. Chen, P. Zhu, and W. Tong, "Integrated sensing and communication in 6g: motivations, use cases, requirements, challenges and future directions," in *1st IEEE International Online Symposium on JC&S*, 23-24 February, 2021.

[3] O. Li et al., "Integrated sensing and communication in 6G: a prototype of high resolution THz sensing on portable device," in *European Conference on Networks and Communications (EuCNC)*, 8-11 June, 2021.

[4] K. Witrisal et al., "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," in *IEEE Signal Processing Magazine*, vol. 33, no. 2, pp. 59-70, March 2016, doi: 10.1109/MSP.2015.2504328.

[5] F. Talebi and T. Pratt, "Channel sounding and parameter estimation for a wideband correlation-based MIMO model," in *IEEE Transactions on Vehicular Technology*, vol. 65, no. 2, pp. 499-508, Feb. 2016, doi: 10.1109/TVT.2015.2404571.

[6] N. Ben Rejeb, I. Bousnina, M. B. Ben Salah, and A. Samet, "Channel parameters estimation using cross-correlation matrix of a wireless SIMO system," *Fourth International Conference on Communications and Networking, ComNet-2014*, 2014, pp. 1-5, doi: 10.1109/ComNet.2014.6840915.

[7] L. Wei, Q. Li, and G. Wu, "Direction of arrival estimation with uniform planar array," *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*, 2017, pp. 1-5, doi: 10.1109/VTCFall.2017.8287882.

[8] F. Wen, N. Garcia, J. Kulmer, K. Witrisal, and H. Wymeersch, "Tensor decomposition based beamspace ESPRIT for millimeter wave MIMO channel estimation," *2018 IEEE Global Communications Conference (GLOBECOM)*, 2018, pp. 1-7, doi: 10.1109/GLOCOM.2018.8647176.

[9] Pei Chen and H. Kobayashi, "Maximum likelihood channel estimation and signal detection for OFDM systems," *2002 IEEE International Conference on Communications. Conference Proceedings*. ICC 2002 (Cat. No.02CH37333), 2002, pp. 1640-1645 vol.3, doi: 10.1109/ICC.2002.997127.

[10] R. Carvajal, J. C. Aguero, B. I. Godoy, and G. C. Goodwin, "EM-based maximum-likelihood channel estimation in multicarrier systems with phase distortion," in *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 152-160, Jan. 2013, doi: 10.1109/TVT.2012.2217361.

[11] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477-489, April 1988, doi: 10.1109/29.1552.

[12] Y. Bar-Shalom, F. Daum, and J. Huang, "The probabilistic data association filter," in *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 82-100, Dec. 2009, doi: 10.1109/MCS.2009.934469.

[13] Paul Meissner, Christoph Steiner, and Klaus Witrisal, "UWB positioning with virtual anchors and floor plan information," *2010 7th Workshop on Positioning, Navigation and Communication*, IEEE, 2010.

[14] Paul Meissner, Thomas Gigl, and Klaus Witrisal, "UWB sequential Monte Carlo positioning using virtual anchors," *2010 International Conference on Indoor Positioning and Indoor Navigation*, IEEE, 2010.

[15] Paul Meissner *et al*., "Analysis of an indoor UWB channel for multipath-aided localization," *2011 IEEE International Conference on Ultra-Wideband (ICUWB)*, IEEE, 2011.

[16] Deissler, Tobias, and Jörn Thielecke., "UWB SLAM with rao-blackwellized Monte Carlo data association," *2010 International Conference on Indoor Positioning and Indoor Navigation*, IEEE, 2010.

[17] Paul Meissner and Klaus Witrisal, "Multipath-assisted single-anchor indoor localization in an office environment," *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, IEEE, 2012.

[18] C. B. Barneto, T. Riihonen, M. Turunen, L. Anttila, M. Fleischer, K. Stadius, J. Ryynänen, and M. Valkama, "Full-duplex OFDM radar with LTE and 5G NR waveforms: challenges, solutions, and measurements," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10, pp. 4042-4054, Oct. 2019.

[19] K. Siddiq, R. J. Watson, S. R. Pennock, P. Avery, R. Poulton, and B. Dakin-Norris, "Phase noise analysis in FMCW radar systems," *2015 European Radar Conference (EuRAD)*, Paris, pp. 501-504, 2015.

[20] C. Wang, Y. Li, Z. Li, K. Zeng, J. He, and G. Wang, "A 3D imaging method for future communication and imaging integrated terminal," *The 2021 CIE International Conference on Radar*.

[21] X. Li, J. He, Z. Yu, G. Wang, and P. Zhu, "Integrated sensing and communication in 6G: the deterministic channel models for THz imaging," *2021 IEEE 32nd annual international symposium on personal, indoor and mobile radio communications (PIMRC)*, 2021, pp. 1-6, doi: 10.1109/PIMRC50174.2021.9569384.

[22] T. Jin, X. Qiu, D. Hu, and C. Ding, "Unambiguous imaging of static scenes and moving targets with the first Chinese dual-channel spaceborne sar sensor," Sensors 17(8), 1709 (2017).

[23] A. Vertiy and S. Gavrilov, "Near-field millimeter wave and microwave tomography imaging," *Proc. Int. Kharkov Symp. Phys. Engrg. Millim. Sub-Millim. Waves (MSMW)*, Jun. 2007, pp. 104-108.

[24] B. Paul, A. R. Chiriyath, and D. W. Bliss, "Survey of RF communications and sensing convergence research," IEEE Access, pp. 252 - 270, 2016.

[25] M. Braun, C. Strum, and F. K. Jondral, "Maximum likelihood speed and distance estimation for OFDM radar," in Proc. *2010 IEEE Radar Conf.*, Washington, DC, May 2010.

[26] G. N. Saddik, R. S. Singh, and E. R. Brown, "Ultra-wideband multifunctional communications/radar system," *IEEE Trans. Microw. Theory Tech.*, pp. 1431-1437, July 2007.

[27] K. Wu and L. Han, "Joint wireless communication and radar sensing systems - state of the art and future prospect," *IET Microwaves Antennas & Propagation*, vol. 7, no. 11, pp. 876-885, 2013.

[28] C. B. Barneto, T. Riihonen, M. Turunen, L. Anttila, M. Fleischer, K. Stadius, J. Ryynänen, and M. Valkama, "Full-duplex OFDM radar with LTE and 5G NR waveforms: challenges, solutions, and measurements," in *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 10, pp. 4042-4054, Oct. 2019.

[29] K. Siddiq, R. J. Watson, S. R. Pennock, P. Avery, R. Poulton, and B. Dakin-Norris, "Phase noise analysis in FMCW radar systems," *2015 European Radar Conference (EuRAD)*, Paris, pp. 501-504, 2015.

[30] C. Jiao, Z. Zhang, C. Zhong, and Z. Feng, "An indoor mmWave joint radar and communication system with active channel perception," in *2018 IEEE International Conference on Communications (ICC)*, Kansas City, MO, 2018.

[31] N. González-Prelcic, R. Méndez-Rial, and R. W. Heath, "Radar aided beam alignment in mmWave V2I communications supporting antenna diversity," in *Information Theory and Applications Workshop (ITA)*, La Jolla, CA, 2016.

[32] Z. Li, S. Yang, and T. Clessienne, "Exploiting location information to enhance throughput in downlink V2I systems," in *2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, United Arab Emirates, 2018.