IEEE Open Journal of
Vehicular Technology

Received 13 October 2021; accepted 29 October 2021. Date of publication 4 November 2021;
date of current version 7 December 2021. The review of this article was arranged by Editor S. Zhang.

*Digital Object Identifier 10.1109/OJVT.2021.3124787*

# Design and PoC Implementation of Mmwave-Based Offloading-Enabled UAV Surveillance System

**TAO YU[1] (Member, IEEE), YOSHITAKA TAKAKU[1], YOHEI KAIEDA[2],
AND KEI SAKAGUCHI[1] (Senior Member, IEEE)**

[1]Tokyo Institute of Technology, Tokyo 152-8550, Japan
[2]Intelligent Systems Laboratory, SECOM Co., Ltd, Tokyo 181-8528, Japan

CORRESPONDING AUTHOR: TAO YU (e-mail: yutao@mobile.ee.titech.ac.jp).

This work was supported in part by the Ministry of Internal affairs and Communications (MIC) of Japan under the Grant 0155-0083,
and in part by SECOM Co., Ltd.

---

**ABSTRACT** Among UAV applications, real-time high-resolution video surveillance systems using unmanned aerial vehicles (UAVs) especially gain our research interests, because high mobilities of UAVs and high qualities of videos can dramatically improve the surveillance performance and evolve the system, especially for artificial intelligence (AI)-based monitoring systems. However, it is challenging to transmit high-resolution videos to ground station in real-time and also impossible to perform AI detection based on these videos on UAVs. To address the challenges, we propose to take full advantages of mmWave communications and to apply it into UAV-based surveillance systems to transmit ultra-high-quality and ultra-high-resolution uncompressed 4 K videos. To increase energy efficiency of UAVs, we propose to offload all computations via mmWave from UAVs to ground station. Moreover, in addition to design of system architecture, proof-of-concept (PoC) prototype hardware of the proposed system is practically developed to demonstrate technical feasibilities of mmWave communication on UAVs and the enhancements to practical surveillance applications. The practical issues in the implementation and the necessary hardware/software applied in the prototype system are explained in detail in this paper. An outdoor PoC experiment is conducted, and the experimental results are consistent with design, so that the effectiveness of the system is confirmed.

**INDEX TERMS** mmWave, UAV, proof-of-concept, surveillance, 4 K uncompressed video, offloading.

---

## I. INTRODUCTION

Thanks to the increasing miniaturization of electronic components and development of high-performance control algorithms in the recent decades, unmanned aerial vehicles (UAVs), so-called drones, have evolved greatly to be more and more operable, functional, productive and affordable than ever before. They are attracting growing interests and popularities from the fields of academia and industries as well as from consumers, because UAVs flexibly and unprecedentedly extend their mobilities from 2D-space to 3D-space, comparing with the conventional ground mobile robots or fixed-location robots. Such unique feature makes them into an ideal platform for varieties of complicated tasks, such as remote sensing, smart agriculture, precise target tracking, surveillance, fast-response disaster monitoring and logistics [1], which were difficult to be accomplished in the past due to the mobility limitations. UAVs are expected to have a global market of 127 billion USD by 2027 [2].

In the wide range of applications which will be deeply revolutionized by UAV-based solutions, real-time high-resolution video surveillance systems using UAVs especially gain our interests. The widely used conventional video surveillance systems have to employ surveillance cameras installed in fixed locations which are pre-configured according to surveillance requirements and scenario, because the cameras are lacking in mobilities and it is difficult to efficiently transmit surveillance video to surveillance center from everywhere. In consequence, the conventional video surveillance systems always severely suffer from cameras' blind spots, owing to the inappropriate camera location selections and the practically inevitable view blockages caused unintentionally by e.g. human and vehicles, or even intentionally by criminals. To address the problem of the limited surveillance coverage of conventional surveillance systems, the performance of surveillance systems can be drastically improved by using UAVs equipped with cameras, sensors and communication modules. UAVs' high mobilities make it possible to flexibly avoid any blockage in view paths, and even to actively track surveillance targets such as intruders. Because of above-mentioned advantages, UAVs have been applied in ever-large numbers of aerial surveillance systems such as border and urban monitoring [3]–[5].

With the development of deep learning and computer vision techniques, full- or semi-artificial intelligence (AI)-based surveillance solutions are being employed in modern surveillance systems, such as AI-assisted aerial monitoring system for firefighting [6], hardhat detection system for construction workers based on surveillance videos by deep learning [7], and urban surveillance system for vehicle re-identification and tracking [8]. Compared with manual inspections of huge amounts of surveillance video by human security men, the AI-based surveillance is less costly and more efficient. One of the influencing factors to performance of AI-based surveillance is the specification of images and videos. Studies have compared effects of different compression algorithms and ratios on recognition accuracies, and the results show that compression could deteriorate accuracies, especially when large compression ratios are used [9], [10]. Moreover, compression also introduces latencies, which is harmful for some surveillance systems which requires early warnings and responses.

The high energy consumption of AI is also an important issue especially for UAV-based surveillances. Experimental evaluation reports that even processing of low-resolution (224×224) images by the state-of-the-art neural network models can consume power as high as tens to hundreds of watts [11]. Obviously, even higher power would be consumed by higher specification images and videos, which however can result in high recognition accuracies. Mobile edge computing (MEC) is a promising solution to improve computation capability and energy stability of mobile devices such as UAVs [12]. It offloads computations from devices to MEC servers with strategies jointly optimized by flight, communication and offload [13], [14].

Therefore, a high-performance surveillance system is expected to be realized by taking advantages of UAV, AI-based detection and MEC. However, to practically implement such combination, there are still several challenges, of which a crucial one is air-to-ground communication. A high-performance air-to-ground communication system can bridge the UAVs and MEC servers, and enable the offloading of high or ultra-high specification videos, e.g., 4 K video, which contains rich information about surveillance targets and can consequently result in better surveillance performance. In micro-wave band, a wide range of UAV communication systems based on e.g. IEEE 802.11 families [15], LTE [16], [17] and 5G [18] has been studied. More detailed literature review about UAV-aided wireless communications, including network architectures, channel characteristics and new opportunities, is given in [19]. Compared with the terrestrial communications, aerial environments lack obstacles. Such characteristic makes the attenuation of interference very slow, which greatly constraints the channel capacity. Recently, the millimeter-wave (mmWave) band, which is mainly used for short-range or line-of-sight (LOS) communication in terrestrial system, has also been investigated for UAV communications because of its capability for ultra-high-speed and ultra-low-latency communications [20], [21].

Research efforts have been made for wireless transmission of video from multiple UAVs. To increase the spectrum efficiency, in [22], an in-band full-duplex multi-UAV communication system with for video surveillance is developed, in which both ground station and UAVs are equipped with high-gain directional antennas. UAV placement and routes optimization algorithms were proposed to provide reliable, real-time, and feasible video transmission in a heterogeneous communication environment in a smart city in [23]. A cooperative UAV video transmission scheme is proposed in [24] by energy-efficient routes and UAV replacement. In [25], a UAV-based secure transmission for scalable videos in hyper-dense networks via caching with small pressure of wireless backhaul is studied.

Other than studies of theoretical exploration, experimental investigations were also conducted on several UAV testbeds to evaluate practical performances. In [26], a testbed for UAV-to-car communication was built, and the impacts of position and antenna orientation/location on performance were measured at 5 GHz band using IEEE 802.11. In [27], an indoor beam-tracking testbed was built for connected robot and UAV over mmWave communications. A prototype of UAV controlled over LTE was demonstrated, and feasibilities of the existing LTE as UAV communication infrastructure were evaluated [28]. In [29], field measurements are conducted for UAVs connected to commercial LTE networks, and the performance of massive UAV deployment is analyzed through simulations. A comprehensive survey on prototypes and experiments of UAV communications can be found in [30]. There is still a lack of experimental demonstration of mmWave-based UAV communication systems, let alone the computation offloading via mmWave aerial communication systems and its benefits to practical applications.

Moreover, existing works mainly contribute to deployments of UAV communication or single technique issue. However, a systematical construction and implementation of completed and practical applications, which are enabled by high-performance UAV communication and related novel techniques, are still absent, because it is complex and costly to establish a testbed which incorporates all facilities such as UAVs, high-performance UAV communication systems,

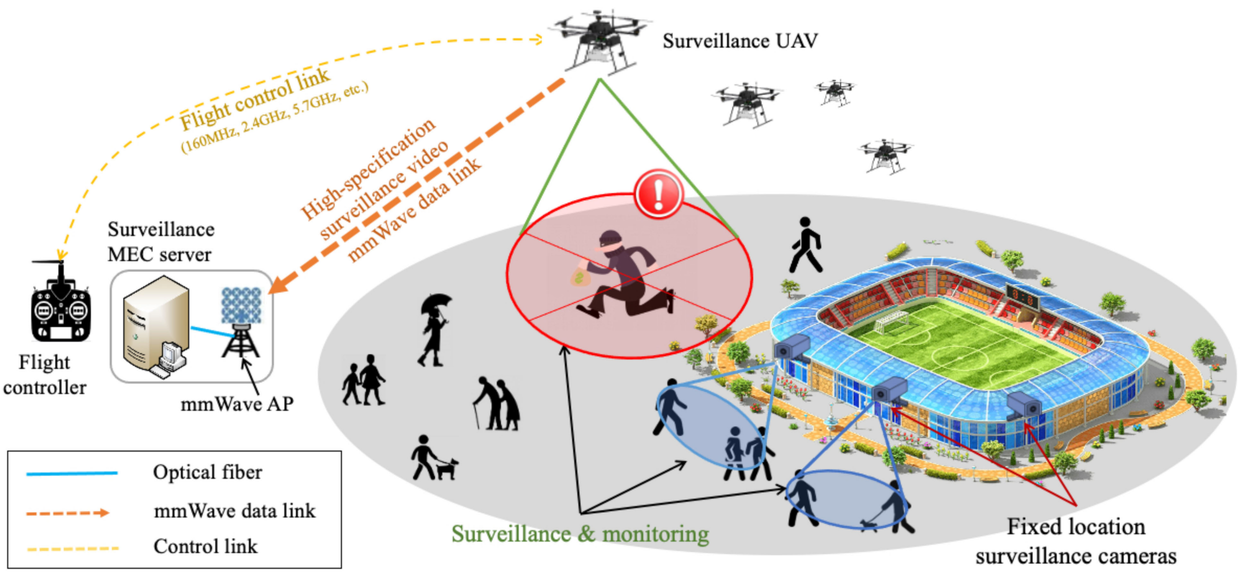


**FIGURE 1.** UAV surveillance system overview.

ground stations and application software. Besides, appropriate experimental fields for UAV flight and wireless radio emitters are also difficult to prepare in countries with strict regulations such as Japan and Europe. In this work, we focus more on the benefits from the well-designed technical combinations of a series of emerging state-of-the-art technologies (i.e. UAV, AI, mmWave, computation-offloading, and high-specification videos) on specific end-use UAV-based applications, i.e. surveillance. The contribution of this work is twofold: 1) the system architecture design of mmWave-based offloading-enabled UAV surveillance system, and 2) the hardware development and proof-of-concept (PoC) implementation.

First, to address the challenges which hinder the progress of UAV-based surveillance applications, in this paper we propose a complete and practical system architecture for a UAV-based surveillance system. The features of the system are as follows:

- To address the performance degradation caused by low-specification surveillance videos, ultra-high-resolution and ultra-high-quality video footages, i.e. uncompressed 4 K raw video, are transmitted from surveillance UAVs to a surveillance center on the ground in real-time for high-accuracy surveillances.
- MmWave-based air-to-ground communication, as the only feasible and qualified candidate of wireless communication techniques, is employed to transmit the huge amount of data of high-specification surveillance videos in real-time from UAVs to ground station.
- Edge computing is introduced to this UAV-based surveillance system. Instead of local processing, all computations are offloaded from UAVs to ground station via mmWave communication, in order to mitigate the energy consumption on UAVs.

Second but more importantly, in addition to the system design and analysis, we focus more on practical development and implementation issues of the PoC hardware prototype to demonstrate technical feasibilities of mmWave communications on UAVs and the consequent enhancements to practical surveillance applications. A PoC testbed is built by practical hardware, and the functionalities are implemented upon developer-friendly open-source software platforms. The implementation considerations and the specific software/hardware applied in the prototype are explained in this paper. These informative details can be a good reference for the readers who plan to practically implement similar hardware prototypes, but to the best of our knowledge, these have not been discussed and disclosed in the previous research in literatures. Specifically, this work is also the first pioneering project of a system-level hardware implementation of a complete system which validates and demonstrates high-mobility UAVs, ultra-high-specification raw video, mmWave air-to-ground communications and computation offloading in the surveillance applications. The PoC experiments are also conducted in an outdoor UAV field to validate whether the developed prototype well reflects the system features and agrees with the design expectations.

The rest of this paper is organized as follows. Sect. II presents the system model and explains the general concept of the proposed system. Configurations and structures of the proof-of-concept prototype are presented in detail in Sect. III. Sect. IV presents the validation experiments results. Finally, Sect. V concludes this paper.

## II. SYSTEM DESIGN

In this section, the concept, architecture and design considerations of the proposed mmWave-based offloading-enabled UAV surveillance system are given. To combine it with the practice, a high-accuracy real-time human recognition system, as one of the most typical and important surveillance applications, is also proposed.

### A. SYSTEM CONCEPT

In Fig. 1, the overview and concept of the proposed mmWave-based offloading-enabled UAV surveillance system is illustrated by one of its typical use-cases, a surveillance system for mega events in stadiums, such as Olympics and concert. In this representative application scene, besides conventional fixed location cameras, surveillance UAVs equipped with mmWave communication modules and high-specification cameras, i.e. 4 K uncompressed video camera, are also deployed to cooperatively fly around the target area and take the surveillance videos. The high mobility of UAVs compensates and eliminates blind spots of fixed location cameras, which makes surveillance much more active and flexible. Moreover, the high-resolution and high-quality surveillance videos contain much richer information about surveillance area and targets than the currently widely used low-specifications videos, and it will result in more accurate surveillance detections.

To reduce complexities of the functional devices on UAVs and efficiently expand the working time of UAVs, all the processing and computations for high-specification videos are offloaded to a MEC server on the ground through mmWave air-to-ground communications for real-time AI-based surveillance analyses such as recognition, alerting and tracking. Because mmWave signals are vulnerable compared with microwave band signals, surveillance actions of UAVs (e.g. flight heights, orientations and trajectories) should be well designed and jointly optimized with mmWave air-to-ground communication to ensure stable mmWave-based video transmission from every surveillance UAVs to ground stations. The key components of the proposed system will be explained in detail as follows.

*1) Surveillance UAVs:* Surveillance UAVs function as the core role in the proposed system. It is the platform which introduces 3D-space mobilities to the system and carries all surveillance/communication hardware in the air. It can take high-specification surveillance videos by on-board sensors (e.g. cameras and LIDARs), which is the input information of surveillance systems. To perform efficient surveillance data transmission, surveillance actions of UAVs (e.g. flight heights, orientations and trajectories) are well designed subjecting to the requirements from mmWave air-to-ground communication. In PoC implementation, medium- or large-sized UAVs are preferable, because the total weight of the prototype hardware can be up to several kilograms. For example, the weight can be up to around 2 kg including mmWave communication module, antennas, a UAV PC, a camera and a gimbal in this work.

One of the motivations and ideologies of this system is to enable ultra-high-specification video transmissions for surveillance UAVs to validate the benefits on applications. For conventional video transmission systems, video encoding/compression, such as M-JPEG, H.264 and H.265 [31], has to be conducted due to the large size of raw video footages. Taking 4 K uncompressed raw video (chroma sub-sampling 4:2:2) as example, the required data rate is 0.2 Gbps per frame, i.e. 12 Gbps, 6 Gbps, 3 Gbps, 1.5 Gbps at 60 fps, 30 fps, 15 fps, 7.5 fps, which are the commonly used frame rates of surveillance cameras. By removing the spatial, temporal visual and coding redundancies in raw video footages, compression rates as high as up to 1000:1 in H.264/H.265 can be obtained. Such bandwidth savings by compressions are in costs of low video quality, high end-to-end latency and high system complexity, which, however, are the crucial factors for UAV-based surveillance systems. High video quality is important for accurate recognitions and interpretations of surveillance targets; low latency is important for early warnings and responses to intruders and threats; and low system complexity is important for weight/size- and energy-constrained UAV-based system implementations. Therefore, in the proposed system, 4 K uncompressed video will be directly transmitted from UAVs to ground station via mmWave communication, and UAVs will not perform any pre-processing or processing on raw videos.

*2) Surveillance MEC Servers:* In the proposed structure, ground stations can be either dedicated MEC servers only for surveillance purpose, or universal MEC servers equipped with mmWave communication modules. Surveillance MEC servers feature powerful computation capabilities and ultra-high-speed mmWave communications. Based on high performance computation enabled by high-end CPUs and GPUs, they can perform real-time and complicated AI-based surveillance processing and detection with high accuracy. This is impossible to be completed on UAVs due to the strong limitation in energy and computation capability. In this work, MEC servers receive high-specification videos from UAVs through mmWave-based data-plane, and then perform real-time AI-based surveillance processing, such as recognition, alerting and tracking. They are supposed to result in low latency and early response to abnormalities and dangers.

In this PoC project, because only one UAV and one MEC server are practically deployed, it is assumed that all computations from the UAV will be offloaded, and MEC server will always have enough resource for UAV. Specially, in this work, an AI-based face recognition application [32], [33] is implemented in MEC server. It is used as an example of the practically and widely used surveillance applications to evaluate and demonstrate the performance improvement by the proposed system. The uncompressed 4 K video received from UAV is used as video source.

*3) Controller:* Flight controller controls flight trajectories or placements of UAVs through control links featuring wide coverage and robust communication. In the PoC project and validation experiment, a manual operator controls UAVs according to path plans which are designed in advance to achieve stable mmWave air-to-ground communication in the experiment field.

However, in desired systems, automatic and optimized flight controls should be performed to guarantee the system performance in complex and dynamic environments. The design and planning of UAVs' placements and flight trajectories as well as communication parameters (e.g. beam selection and channel assignment and power allocation) are always formulated into a joint optimization problem to maximize system performance indicators such as communication throughputs, battery-life, security, and task execution efficiency [34]-[36]. For the development and implementation of automatic UAV control systems instead of the manual operations, the Robot Operating System (ROS) [37], which is a general-purpose framework for robotics development, and the DJI Software Development Kit (DJI SDK) [38], which is a dedicated toolkit for the UAV application development of DJI UAVs, are the two most widely used solutions depending on the model of UAVs.

In addition, practical systems should have more complicated structures (e.g. multiple UAVs/MEC servers, dynamic network topology) and functions (e.g. the cooperative operations, MEC container migration). Therefore, effective MEC resource managements must also be performed to reduce the potential resource collisions which decrease system performance. In MEC-enabled UAV systems, MEC resources (e.g. computation capability, UAV association) and UAVs' tasks offloading planning (e.g. partial or binary) should be also jointly optimized with the above-mentioned flight trajectory and communication parameters [12]–[14]. For practical implementation and management of MEC servers, virtualization and container tools, such as Docker [39] and Kubernetes [40], are widely applied.

### B. COMMUNICATION ARCHITECTURE

In order to guarantee effective and efficient video transmission for surveillance in the proposed system, both of advanced mmWave communication and micro-wave band communication techniques are employed in the system, and they play different roles in the system.

*1) High-Speed Air-to-Ground Access Link:* High-specification surveillance videos, containing rich real-time information about surveillance area and targets, are crucial for the performance of surveillance systems.

As explained in Sect.II.A.1, transmission for uncompressed 4 K raw video requires throughputs as large as from 1.5 Gbps (at 7.5 fps) to 12 Gbps (at 60 fps). In microwave bands, the conventional aerial communication systems such as those reported in [15], [16] and [18], is impossible to provide air-to-ground communications meeting these requirements. The cutting-edge standardized IEEE 802.11ax can support maximum physical (PHY)-layer data rate of 1.2 Gbps (8-channel bonding, 160 MHz bandwidth, single stream, modulation and coding scheme (MCS)-11) which typically results in transmission control protocol (TCP)-layer data rate of 0.4 Gbps $\sim$ 0.6 Gbps [41]. (It is noted that in air-to-ground propagation environment, MIMO transmission cannot be established because of lack of multi-path reflections.) On the other hand, for mmWave-based techniques, the standardized IEEE 802.11ad supports maximum PHY data rate of 6.8 Gbps (single channel, 2.16 GHz bandwidth, single stream, MCS-24) which typically results in TCP data rate of about 2.3 Gbps $\sim$ 3.4 Gbps [42]; the cutting-edge IEEE 802.11ay can support up to 38 Gbps (4-channel bonding, 8.64 GHz bandwidth, single steam, MCS-19) which typically results in TCP data rate of about 12 Gbps $\sim$ 19 Gbps [43]. Therefore, to enable ultra-low latency transmission of uncompressed 4 K, even 8 K, videos, the only feasible solution is mmWave-based communication.

To address the large propagation attenuation in mmWave band, directional antennas of narrow beam width are necessary. Moreover, in the target surveillance use-case (e.g. surveillances for mega events in stadium), because the distance between UAVs and ground stations is not long (typically within hundreds of meters), alignments of antenna beams should be fast and wide enough in response to fast-speed motions of UAVs. To achieve the optimized communication and computation resource management in the proposed system, air-to-ground link need to be well designed and optimized based on real-time context information of the surveillance UAVs and ground stations.

Another important issue in mmWave communication systems is the blockage effect due to the short wavelength, which greatly hinders the applications of mmWave in terrestrial communication systems. However, in aerial systems, the combination of UAV and mmWave communication can introduce the capability of mobility to mmWave transceivers and complement each other's advantages. Beam controls and flight controls of UAV can be jointly performed to guarantee LOS UAV-to-ground path. Moreover, because the proposed system is expected to be used in security surveillances of stadium and parks, blockages between UAVs and ground station could be rare if position of ground station is well selected, e.g. on the top of buildings.

In the case of multi-UAV system with dense deployment or shared spectrum with other system, co-channel interference (CCI) must be taken into considerations due to lack of obstacles, especially in UAV-to-UAV channels. Because highly directional steerable antennas are used in mmWave systems, besides the radio resource management schemes such as power control, adaptive MCS, and dynamic spectrum/time resource allocation which are widely used in lower frequency systems, interference elimination based on beam management and spatial reuse are also considered as promising solutions in mmWave systems, such as dynamic temporal-spatial reuse of spectrum for mmWave in [44], and deep-learning-based beam management and interference coordination for dense mmWave network in [45]. Moreover, such methods are especially fit for multi-UAV system because of the unique features of multi-UAV system, i.e., mobility, controllability, and cooperativity. Therefore, in practical implementation of multi-UAV mmWave system, conventional interference cancellation and resource management schemes combined with cooperative trajectory control of multiple UAVs need to be considered.

*2) Ubiquitous and Reliable Control Link:* In the proposed system, the control link is to guarantee safe flying and stable video transmission. In this PoC project, it is responsible to send flight control information and collect UAV's flight status information. Those messages carry information which is crucial for system performance and safety flight, and they need to be updated continuously and frequently.

To meet these requirements, the control link is expected to have wide coverage and reliability but does not necessitate the data rate as high as data link. Therefore, the separation architecture of data and control links is adopted. The existing conventional communication techniques in microwave bands, such as 5 G sub-6 GHz, LTE, Wi-Fi and WiMAX systems, are qualified to provide control signals with wide coverage and reliability. For example, the LTE and WiMAX which have very high penetration and coverage in both urban and rural areas, and Wi-Fi systems, which are widely used in daily life and very convenient for deployment, are good candidates for control-plane access technologies. Integration with such existing infrastructures will be an economical and practical solution.

## III. PROOF-OF-CONCEPT PROTOTYPE

In order to validate the benefits from proposed system architecture on specific surveillance applications, a PoC prototype system is practically developed and implemented based on a series of state-of-the-art hardware and open-source software. The prototype system is built to fill the gap between practical developments and the theoretic explorations in plentiful literatures. It aims to demonstrate the feasibility of mmWave communication techniques in the aerial communication systems, and to address the issues in the prototype development and implementation with practical software/hardware. More specifically, in the prototype, a UAV-based video surveillance system offloads all the computation of surveillance detections by transmitting uncompressed 4 K video to a ground station in real-time. Details of software/hardware model, parameters, and configurations in the implementation are presented in this section.

Without loss of the key features of the system, the testbed is built as minimum structure as shown in Fig. 2. The PoC system consists of one surveillance UAV, one ground station (MEC server) and one controller. The communication architecture includes one mmWave data link and control links. Despite of a minimum system structure, the PoC system is fully functional and we believe it is able to demonstrate all system features as explained in the previous sections. Moreover, this PoC system is scalable and is ready for extensions to multiple surveillance UAVs and ground servers. The detailed system diagram in Fig. 3 shows the structure of hardware components of the surveillance UAV, ground server and controller in the prototype with communication connections of control-plane and data-plane. The UAV PC, MEC PC, 4 K camera, and mmWave modules (WiGig) are all products chosen from the market. The gimbal and mmWave lens antenna are custom-made based on technical requirements according to the system design. The UAV is refitted to a surveillance platform based on a commercial model. The detailed specifications and related information of the hardware components employed in this PoC implementation are summarized in Table 1, which can be a reference for the readers who plan to build a similar system.

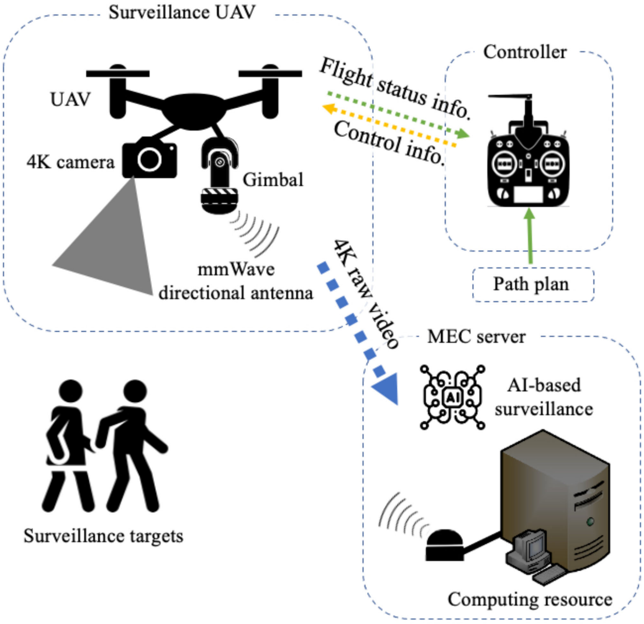


**FIGURE 2.** Illustration of the system architecture of the PoC implementation.

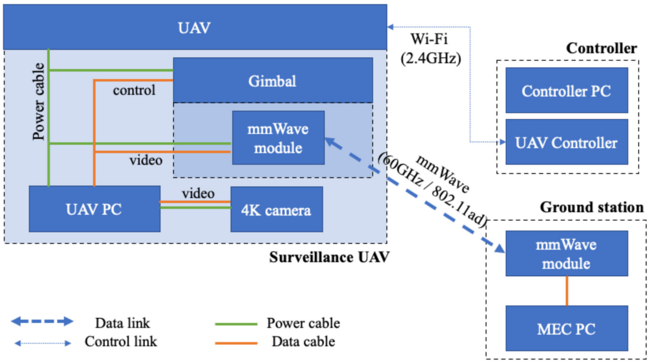


**FIGURE 3.** System diagram of the PoC implementation.

### A. MMWAVE AIR-TO-GROUND COMMUNICATION MODULE

In the data-plane of this work, the certified IEEE 802.11ad (WiGig) devices in the license free 60 GHz mmWave band are adopted [42]. The high-directivity antennas are necessary for the air-to-ground mmWave communication because of the large propagation loss in this band. The 2 × 8 antenna array with high-directivity lens antenna and its radiation patterns are shown in Fig. 4. The lens antennas are employed to enhance the coverage and enable mmWave communication up to several hundreds of meters. The gain is 25.4 dBi, and beam width is 6.4° (horizontal) and 2.9° (vertical). It is lightweighted (total weight around 500 g) compared with the payload, and thus can be equipped on UAVs without potential safety risks. The scanning angle range of the antenna is ±13.5° (horizontal) and ±7° (vertical).

The high-performance beamforming must be performed due to the highly directional nature of mmWave, the UAVs'

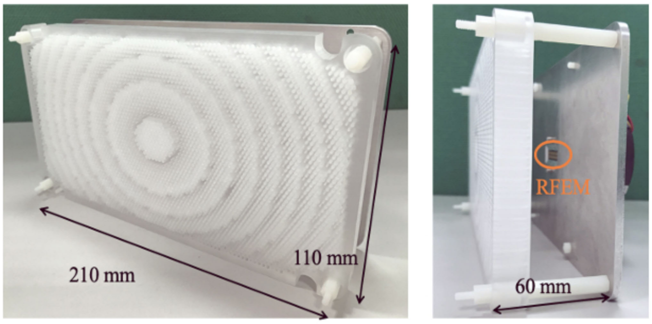
**TABLE 1.** **System Parameters**

| Hardware | Specifications |
|---|---|
| mmWave module | Module: Intel 18260NGW<br>Protocol: IEEE 802.11ad (WiGig)<br>Maximum MCS: 12 |
| mmWave antenna | Module: Intel lens antenna<br>Scan Angle: ± 13.5 ° (Hor.), ± 7 ° (Ver.)<br>HPBW: ± 6.4 ° (Hor.), ± 2.9 ° (Ver.)<br>Gain: 25.4 dBi<br>Weight: 500 g |
| Surveillance UAV | Module: ACSL-PF1<br>Size: 117 cm × 106 cm × 41.8 cm<br>Weight: 6.4 kg<br>Flight time: 45 min<br>Payload: 3 kg |
| Camera | Model: See3CAM_130<br>Size: 80 mm × 15 mm × 9.7 mm<br>Resolution: 3840 × 2160, 1920 × 1080,<br>Pixel Format: UYVY<br>Power Consumption: 1.85W<br>Interface: USB 3.1 Gen1 |
| UAV onboard PC | Model: Intel NUC<br>OS: Ubuntu 16.04<br>Wireless LAN: IEEE 802.11n<br>Weight: 520 g |
| Edge PC | OS: Ubuntu 16.04<br>GPU: Nvidia GTX 1080 Ti<br>CPU: Core i7 6900k<br>Memory: 32 GB |

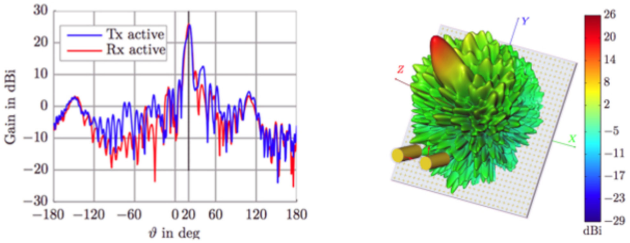
high-speed mobility and the vibrations caused by winds and the mechanical structures of propellers. IEEE 802.11ad defines a selection-based beamforming protocol, in which the switched antenna array is used to find out the best beam pairs in the transmitter and receiver for a WiGig communication link and to realize user association based on the maximum received power [42]. A mechanical rotatable gimbal and stabilizer is also installed on the underside of the UAV to assist the beamforming and the polarization-plane matching.

In the experiment, because the receiver is put on the ground and beam width of the lens antenna is very narrow, multi-path effect of reflections, e.g. from the ground and surrounding buildings, can be ignored, so that air-to-ground channel in this work is approximately free-space propagation.

According to the data sheet of the device and our pre-test, its driver can keep throughput up to more than 1.5 Gbps in TCP layer, i.e. more than 3 Gbps in PHY layer, with MCS-12 in single carrier mode. (Theoretical PHY throughput of MCS-12 is up to 4.62 Gbps [42].) The pre-test is performed on the ground to confirm the performance of mmWave devices. Both transmitter and receiver are directly placed on flat playground to eliminate ground reflection, and their main lobes face to each other. Measurement results are shown in Fig. 5, in which blue solid curve is theoretical PHY throughput, blue dash curve is required throughput, blue dot curve is measured TCP throughput, and red curve is theoretical SNR. It shows that the devices can guarantee TCP throughput larger than required throughput within 100 m. Therefore, the maximum distance from UAV to receiver will be kept smaller than 100 m in experiment.



(a) Photo of the lens antenna



(b) Radiation pattern

**FIGURE 4.** **mmWave lens antenna.**

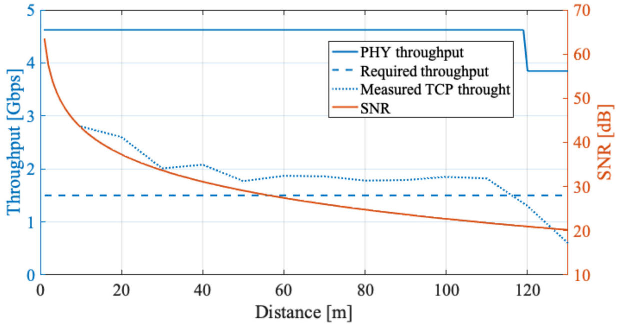


**FIGURE 5.** **Pre-test throughput.**

For terrestrial systems, when channel qualities deteriorate, in order to maintain communication quality, transmitter has to switch to lower MCS index and the consequent lower throughput. In the validation experiment of this PoC prototype, flight trajectories are pre-designed to maintain the above-mentioned required SNR and data rate. In real system, such mechanism is expected as active UAV maneuvering.

### B. SURVEILLANCE UAV

Figure 6 shows the photo and the structure of the UAV employed in this research. It is refitted into a mmWave-enabled dedicated surveillance UAV based on a common commercial model bought from the market. As explained in the previous

**FIGURE 6.** The refitted surveillance UAV.

subsection, joint mechanical and electrical beamforming is adopted in this research. A mechanical 3D rotatable gimbal and stabilizer is installed on the underside of UAV body to enhance stability and scanning range of beamforming and polarization-plane matching. For the compact size and light weight, the UAV PC, mmWave module, array antenna and lens have been integrated together by a plastic cover, and they are attached to the gimbal in the UAV as shown in Fig. 6.

4 K uncompressed video with resolution 3840×2160 is taken by a lightweight 4 K camera attached to UAV, and data rate per frame is around 200 Mbps. Due to the limitation by maximum throughput in TCP layer of current mmWave devices, i.e. up to 1.5 Gbps, frame rate of this camera is set to be 7.5 fps. Please note that the throughput can be increased by using channel bonding or using devices with higher specification supporting higher MCS.

Other advanced sensors (e.g. GPS, altimeter, speed sensor, and gyroscope) are also indispensable in UAV, and sensing data from these sensors are used by central controller to coordinate communication and maneuvering. UAV PC is a lightweight Intel NUC PC, whose functions are to arrange video recording and mmWave connections.

### C. CONTROLLER

In this PoC system, commands sent to UAV and real-time context information collected from surveillance are all through Wi-Fi (2.4 GHz band, IEEE 802.11n). The coverage and capacity of Wi-Fi are sufficient for this research because this PoC system is with minimum system architecture and scale. For larger-scale system, the existing LTE cellular network and WiMAX (2.5 GHz band, IEEE 802.16) are recommended because of their penetration, coverage and stability.

### D. MEC SERVER (GROUND STATION)

In this PoC system, single MEC server is practically deployed on the ground. It is equipped with a WiGig mmWave communication module, which is with the same model and configuration as UAV as shown in Fig. 3, to set up the mmWave data-plane and receive uncompressed 4 K video transmitted from UAV.

The MEC server features large computation capability compared with UAV. It has high-performance GPU (Nvidia GTX 1080 Ti) and CPU (Intel Core i7 6900 k) for the surveillance recognition. Moreover, a GPU-based real-time face detection and recognition application, considered as one of the most widely used surveillance applications, was implemented in MEC server using uncompressed 4 K video sent from UAV.

In this work, to demonstrate the benefits of computation-offloading from UAV to MEC server via mmWave, a full-offloading policy is adopted, with the following two considerations. 1) MmWave aerial communication breaks through communication constraints. 2) Computation resources on the MEC overwhelmingly have advantages compared with energy- and computation-constrained UAVs. However, in systems with constrained resources of communication and MEC, it is necessary to optimize offloading policies and split tasks to balance resources, performance, energy and latency [12]–[14].

## IV. EXPERIMENTS AND RESULTS

A UAV communication experiment is conducted by using the hardware described in the previous sections in order to validate the PoC system and to demonstrate feasibilities of mmWave communication in UAV systems. As a representative example for the usage of high-specification videos in surveillance applications, a face recognition experiment is also conducted to show the performance improved by the mmWave communication. Please note that the video transmission experiment and the surveillance recognition experiment are conducted separately due to the safety considerations and the related regulations/laws in Japan and the TokyoTech. However, since the two experiments use the same hardware configurations, we believe that the results are valid and effective to evaluate the system performance.

### A. HIGH-SPECIFICATION VIDEO TRANSMISSION VIA MMWAVE

In the outdoor experiment of mmWave transmission from UAV, mmWave antenna of the ground station (receiver) is placed very close to the ground. In the experiment field there is no effective radio-wave reflectors (e.g. buildings and mountains) around, so that nearly pure LOS communication is expected, and multi-path effect can be considered to be negligible.

Because there is only one MEC user (the surveillance UAV) in the system, ground server is always ready and idle before

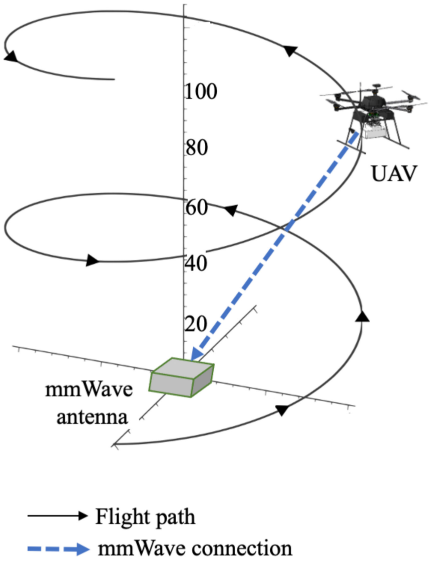
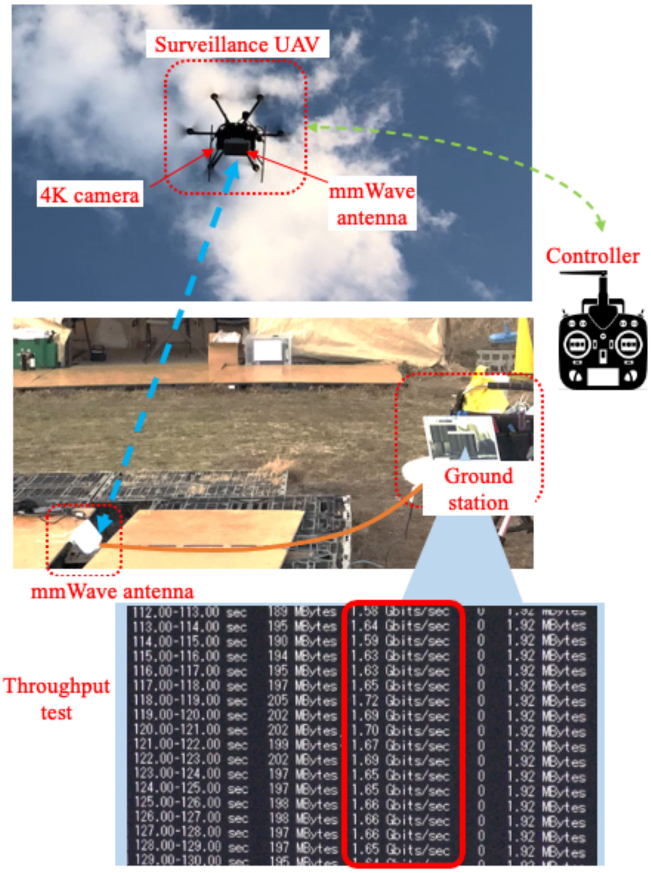


**FIGURE 7.** **Flight path of UAV in experiment.**

data-plane connection is set up. The transmitter and receiver continuously scan and try to establish connection with the best beam ID. As soon as mmWave communication is possible, UAV and ground station APs (Access Points) will setup connection and perform uncompressed 4 K video transmission. In the experiment, surveillance UAV flies over ground station, and ascends higher and higher up to 100 m, which is within effective coverage for 1.5 Gbps throughput according to the calculation in the previous section. Fig. 7 illustrates UAV flight path during experiment. Fig. 8 depicts experiment environment and configuration by photos taken on site.

Firstly, during this climbing process of UAV, real-time TCP layer throughput of mmWave air-to-ground communication from UAV to ground station is measured by the network performance measurement tool, iperf3 [47]. The result confirms that throughputs larger than 1.5 Gbps can be achieved throughout the experiment. A screenshot of throughput test in experiment is also given in Fig. 8. Such result indicates that the data-plane in this PoC system is capable to support transmission of uncompressed 4 K video at 7.5 fps as designed for high-performance surveillance.

The average round trip time (RTT) of the mmWave connection in the experiment, which is twice of its end-to-end (E2E) latency, is measured as average of 2.5 ms and the minimum of 2.1 ms. For comparison, the average RTT of IEEE 802.11n connection, which is widely used in current commercial UAV systems, is measured to be more than 80 ms. E2E latency is approximately equal to sum of processing latency, transmission latency, queuing latency and propagation latency. During them, latency caused by processing, queuing and propagation in the above-mentioned two kind of connections are typically very close. Therefore, the improvement in latency performance mainly benefits from the ultra-low transmission latency of mmWave due to its ultra-high-speed transmission.



```
112.00-113.00 sec   189 MBytes  1.58 Gbits/sec   0   1.92 MBytes
113.00-114.00 sec   195 MBytes  1.64 Gbits/sec   0   1.92 MBytes
114.00-115.00 sec   190 MBytes  1.59 Gbits/sec   0   1.92 MBytes
115.00-116.00 sec   194 MBytes  1.63 Gbits/sec   0   1.92 MBytes
116.00-117.00 sec   195 MBytes  1.63 Gbits/sec   0   1.92 MBytes
117.00-118.00 sec   197 MBytes  1.65 Gbits/sec   0   1.92 MBytes
118.00-119.00 sec   205 MBytes  1.72 Gbits/sec   0   1.92 MBytes
119.00-120.00 sec   202 MBytes  1.69 Gbits/sec   0   1.92 MBytes
120.00-121.00 sec   202 MBytes  1.70 Gbits/sec   0   1.92 MBytes
121.00-122.00 sec   199 MBytes  1.67 Gbits/sec   0   1.92 MBytes
122.00-123.00 sec   202 MBytes  1.69 Gbits/sec   0   1.92 MBytes
123.00-124.00 sec   197 MBytes  1.65 Gbits/sec   0   1.92 MBytes
124.00-125.00 sec   197 MBytes  1.65 Gbits/sec   0   1.92 MBytes
125.00-126.00 sec   198 MBytes  1.66 Gbits/sec   0   1.92 MBytes
126.00-127.00 sec   198 MBytes  1.66 Gbits/sec   0   1.92 MBytes
127.00-128.00 sec   197 MBytes  1.66 Gbits/sec   0   1.92 MBytes
128.00-129.00 sec   197 MBytes  1.65 Gbits/sec   0   1.92 MBytes
129.00-130.00 sec   195
```

**FIGURE 8.** **Photo of experiment.**

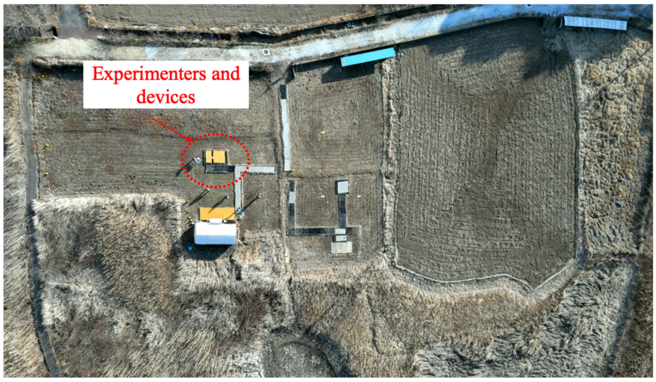


**FIGURE 9.** **A screenshot of the video taken from 100 m height.**

Secondly, UAV repeats the same flight procedure described above, during which experiment of high-specification video transmission through mmWave is conducted. Ultra-high-resolution and ultra-high-quality 4 K raw video is taken by UAV and transmitted to ground station through mmWave connections. The reliable transmission from up to height of 100 m is confirmed to be successful. Fig. 9 gives a snapshot from uncompressed 4 K video which is received at ground station in the experiment. The video of experiment is also available online [1].

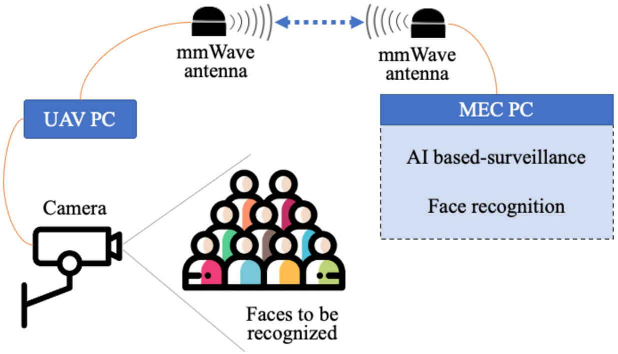
The experiment results successfully verify reliability and the validity of the wireless communication infrastructures for data-plane and control-plane in the proposed system for surveillance. It indicates that mmWave communication is a technically and practically feasible candidate for air-to-ground communication, and it has great potential to enable varieties of UAV-based applications by ultra-high-speed and ultra-low-latency communication. Moreover, higher throughput than 1.5 Gbps achieved in this experiment can be realized by using channel bonding or devices supporting higher MCS.

### B. SURVEILLANCE APPLICATION

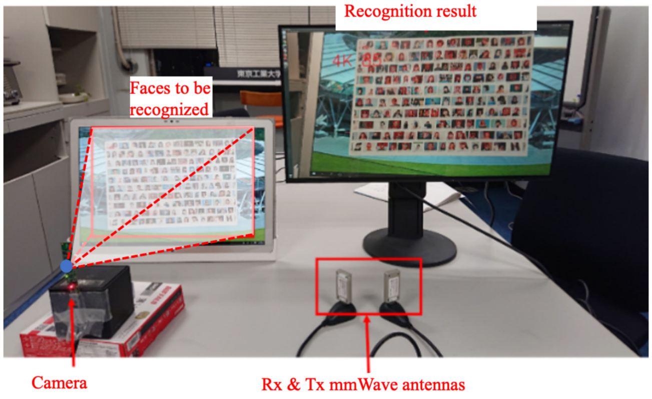
Because of safety considerations and related laws/regulations in Japan and TokyoTech, it is difficult to directly fly the refitted surveillance UAV very close to the people and perform the surveillance detection. Therefore, as a compromise, we conducted a face recognition experiment on ground separately. However, we kept the consistency of hardware and configurations between the outdoor mmWave air-to-ground video transmission experiment and this indoor surveillance recognition experiment. Software from open source community is used for implementation of face recognition application [32], [33] and [46]. Both CPU- and GPU-based recognitions are tested. Fig. 10(a) shows the architecture and configuration of the face recognition experiment. It uses the same camera, PCs and mmWave modules as the outdoor video transmission experiment. It is noted that smaller-sized mmWave antennas are used in this indoor experiment for convenience, but video frame rate and corresponding communication throughput are all kept the same as the outdoor mmWave air-to-ground transmission experiment. Please note that although numerical analysis is given, it is rather a qualitative experiment than a quantitative one.

In this experiment, a 4 K camera takes uncompressed 4 K video, and then UAV PC transmits video stream to MEC PC through mmWave in real-time. By using the videos, face recognitions are performed in ground server. Photo of the experiment setup is given in Fig. 10(b). In order to quantify experiment results, a series of photos of faces are printed on paper and shown in a monitor in the left of Fig. 10(b) to simulate audience in a stadium. Distance between the camera and the monitor is set to make the captured faces have the same size with 26 m away in real world.

The original test video is taken in 4 K resolution (3840×2160) at 7.5 fps in pixel format of uncompressed UYVY (16 bpp). Then, for comparisons, the video is converted into different resolutions and qualities, i.e. compressed 4 K, uncompressed 2 K (1920×1080), and compressed 2 K at the same frame rate. Moderate video compressions are performed to raw videos at 0.1 bpp by a H.264 encoder.



(a) Configuration of face recognition experiment



(b) Photo of experiment setup

**FIGURE 10.** **mmWave lens antenna.**

**TABLE 2.** **Recognition Accuracy**

| Mode | BPP | CPU-based recognition | GPU-based recognition |
|---|---|---|---|
| Comp. 2K | 0.1 | 27% | 50% |
| Uncomp. 2K | 16 | 33% | 52% |
| Comp. 4K | 0.1 | 56% | 86% |
| Uncomp. 4K | 16 | 63% | 92% |

To demonstrate benefits on energy saving owing to computation offloading by mmWave, CPU- and GPU-based recognition are both conducted, and the latter uses deeper and larger networks.

Recognition accuracies with different video specifications and recognizers are shown in Table 2. The setup of compressed 2 K video is used as the baseline for comparison, because it is currently the most widely video specification for real-time video transmission from UAVs. The results confirm the performance improvements thanks to higher video resolution and higher video quality. The system gets great benefits by increasing resolution from 2 K to 4 K in both compressed and uncompressed video, because small face regions in low-resolution videos lack detailed information and make it difficult to detect face landmarks. Uncompressed videos also show performance superior to compressed videos, because

---

[1] https://youtu.be/L7OfRoHZ9jE

**TABLE 3.** **Power Consumption**

| | Idle | Comm. | CPU-based recognition | GPU-based recognition |
|---|---|---|---|---|
| UAV PC | 5.1 W | 9.2 W | 17.3 W | — |
| MEC PC | 75.5 W | 81.6 W | 87.1 W | 353.4 W |

video compression causes the non-reversible distortion and details lost at low bpp (Bits Per Pixel). Moreover, if there are mobilities of people and the inter-frame compressions (e.g. H.264/H.265 inter-modes) are used, it can be reasonably inferred that the benefits owing to uncompressed 4 K video should become larger.

The times consumed by CPU-based (in UAV onboard PC) and GPU-based (in MEC PC) recognitions of 4 K video are measured as average of around 660 ms and 35 ms per frame, respectively. The results indicate that the recognition latency is greatly improved by offloading the heavy computation via mmWave air-to-ground link.

Energy consumptions in different cases are also measured. First, idle powers of MEC and UAV onboard PCs are measured by a power meter. Then, the energy consumed by a functional module of interest can be derived by turning it on and then measuring the increased power. For example, the energy consumption of mmWave communication can be derived by using *iperf* command to make the link full load and then measuring the increased power in both MEC and UAV onboard PCs. Following the same way, the energy consumption of CPU- and GPU-based recognitions can also be measured. Precision of results is improved by averaging the power meter readings of repeated measurements. Table 3 shows the experiments results. The UAV onboard PC's idle energy consumption is 5.1 W, which also includes the energy consumption of the 4 K camera which captures uncompressed 4 K video. When the mmWave communication link is in full load, the power increases to 9.2 W, which implies that communication consumes 4.1 W. When the communication link is down and CPU-based recognition is performed in onboard PC, the power consumption increases to 17.3 W. On the other hand, the MEC PC's idle energy consumption is 75.5 W. When mmWave communication link is set up and in full load, the power increases to 81.6 W. When the mmWave link is up and CPU- or GPU-based recognition is performed in MEC PC, the power consumption increases to 87.1 W or 353.4, respectively. The results show that UAV takes 4.1 W to fully offload computations to MEC though mmWave communication, as a result of which much more accurate recognition can be enabled and up to 271.8 W power can be saved compared with processing locally on UAV.

## V. CONCLUSION

In this work, we propose a complete and practical system architecture for UAV-based surveillance. It features uncompressed 4 K video transmission, mmWave air-to-ground communication and computation offloading. Moreover, we practically develop and implement the PoC hardware prototype of the proposed system, and this is the first pioneering project of a system-level hardware implementation of such a complete UAV system. We believe that it can be a good reference for the readers planning to set up similar UAV testbeds. Validation experiments results agree well with the design expectations. MmWave communication enables UAV high-specification video transmission and shows great potentials in UAV communications systems and the practical applications. The experiment also confirms that in surveillance applications, high-specification video can greatly improve the recognition accuracies. Computation offloading, enabled by mmWave, is proved to be a very effective solution for UAV's energy issue.

## REFERENCES

[1] H. Shakhatreh *et al.*, "Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges," *IEEE Access*, vol. 7, pp. 48572–48634, 2019.

[2] M. Micha, A. Wisniewski, and J. McMillan. "PwC global report on the commercial applications of drone technology," PricewaterhouseCoopers, Tech. Rep., 2016.

[3] C. C. Haddal and J. Gertler, "Homeland security: Unmanned aerial vehicles and border surveillance," Congressional Res. Service, Washington, DC, USA, *Tech. Rep.*, 2010.

[4] N. Chen *et al.*, "Dynamic urban surveillance video stream processing using fog computing," in *Proc. IEEE 2nd Int. Conf. Multimedia Big Data*, 2016, pp. 105–112.

[5] A. Giyenko, and Y. Im Cho, "Intelligent UAV in smart cities using IoT," in *Proc. 16th Int. Conf. Control, Automat. Syst.*, 2016, pp. 207–210.

[6] K. W. Sung *et al.*, "PriMO-5G: Making fire-fighting smarter with immersive videos through 5G," in *Proc. IEEE 2nd 5G World Forum*, 2019, pp. 280–285.

[7] Q. Fang *et al.*, "Detecting non-hardhat-use by a deep learning method from far-field surveillance videos," *Automat. Construction*, vol. 85, pp. 1–9, 2018.

[8] X. Liu, W. Liu, T. Mei, and H. Ma, "A deep learning-based approach to progressive vehicle re-identification for urban surveillance," in Proc. Eur. Conf. Comput. Vis., 2016, pp. 869–884.

[9] K. Delac, M. Grgic, and S. Grgic, "Image compression effects in face recognition systems," in *Face Recognition*. London, U.K.: Intech Open Access, 2007.

[10] A. Mascher-Kampfer, H. Stögner, and A. Uhl, "Comparison of compression algorithms impact on fingerprint and face recognition accuracy," in *Proc. SPIE 6508, Visual Communications and Image Processing*, 2007, Art. no. 650810.

[11] D. Li, X. Chen, M. Becchi and Z. Zong, "Evaluating the energy efficiency of deep convolutional neural networks on CPUs and GPUs," in *Proc. IEEE Int. Conf. Big Data Cloud Comput. (Bdcloud), Social Comput. Netw. (Socialcom), Sustain. Comput. Commun.*, 2016, pp. 477–484.

[12] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surv. Tut.*, vol. 19, no. 4, pp. 2322–2358, Oct.–Dec. 2017.

[13] T. Bai, J. Wang, Y. Ren, and L. Hanzo, "Energy-efficient computation offloading for secure UAV-edge-computing systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6074–6087, Jun. 2019.

[14] M. Messous, A. Arfaoui, A. Alioua, and S. M. Senouci, "A sequential game approach for computation-offloading in an UAV network," in *Proc. IEEE Global Commun. Conf.*, 2017, pp. 1–7.

[15] S. Hayat, E. Yanmaz, and R. Muzaffar, "Survey on unmanned aerial vehicle networks for civil applications: A communications viewpoint," *IEEE Commun. Surv. Tut.*, vol. 18, no. 4, pp. 2624–2661, Oct.–Dec. 2016.

[16] "Enhanced LTE support for aerial vehicles," 3GPP TR 36.777, 2019.

[17] W. Mei and R. Zhang, "Aerial-ground interference mitigation for cellular-connected UAV," *IEEE Trans. Wireless Commun*. vol. 28, no. 1, pp. 167–173, Feb. 2021.

[18] A. Garcia-Rodriguez, G. Geraci, D. Lopez-Perez, L. G. Giordano, M. Ding, and E. Bjornson, “The essential guide to realizing 5G-connected UAVs with massive MIMO,” *IEEE Commun. Mag.*, vol. 57, no. 12, pp. 84–90, Dec. 2019.

[19] Y. Zeng, R. Zhang, and T. J. Lim, “Wireless communications with unmanned aerial vehicles: Opportunities and challenges,” *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.

[20] W. Xia, M. Polese, M. Mezzavilla, G. Loianno, S. Rangan, and M. Zorzi, “Millimeter wave remote UAV control and communications for-Public safety scenarios,” in *Proc. 16th Annual IEEE Int. Conf. Sens., Commun., Netw.*, 2019, pp. 1–7.

[21] W. Khawaja, O. Ozdemir, and I. Guvenc, “UAV air-to-ground channel characterization for mmwave systems,” in *Proc. IEEE 86th Veh. Technol. Conf.*, 2017, pp. 1–5.

[22] T. Yu, S. Imada, K. Araki, and K. Sakaguchi, “Multi-UAV full-duplex communication systems for joint video transmission and flight control,” in *Proc. 11th Annu. Comput. Commun. Workshop Conf.*, 2021, pp. 1423–1428.

[23] J. Yong, Z. Qian, and W. Yang, “UAV cluster-based video surveillance system optimization in heterogeneous communication of smart cities,” *IEEE Access* vol. 8 pp. 55654–55664, 2020.

[24] P. Cumino *et al.*. “Cooperative UAV scheme for enhancing video transmission and global network energy efficiency,” *MDPI Sensors*, vol. 18, no. 12, 2018, Art. no. 4155.

[25] N. Zhao *et al.*, “Caching UAV assisted secure transmission in hyperdense networks based on interference alignment,” *IEEE Trans. Commun.*, vol. 66 no. 5, pp. 2281–2294, May 2018.

[26] S. A. Hadiwardoyo, C. T. Calafate, J. C. Cano, Y. Ji, E. Hernández-Orallo, and P. Manzoni, “Evaluating UAV-to-car communications performance: From testbed to simulation experiments,” in *Proc. 16th IEEE Annu. Consumer Commun. Netw. Conf.*, 2019, pp. 1–6.

[27] K. Heimann, J. Tiemann, D. Yolchyan, and C. Wietfeld, “Experimental 5G mmwave beam tracking testbed for evaluation of vehicular communications,” in *Proc. IEEE 2nd 5G World Forum*, 2019, pp. 382–387.

[28] L. Sundqvist, “Cellular controlled drone experiment: Evaluation of network requirements,” Master thesis, School Electr. Eng., Aalto University, Otaniemi, Espoo, Finland, 2015.

[29] X. Lin *et al.*, “Mobile network-connected drones: Field trials, simulations, and design insights,” *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 115–125, Sep. 2019.

[30] Q. Song, Y. Zeng, J. Xu, and S. Jin, “A survey of prototype and experiment for UAV communications,” *Sci. China Inf. Sci*. vol. 64, 2021, Art. no. 140301.

[31] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, “Overview of the h 264/AVC video coding standard,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[32] D. E. King, “Dlib-ml: A. machine learning toolkit,” *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, 2009.

[33] O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep face recognition,” in *Proc. Brit. Mach. Vis. Conf.*, 2015, pp. 1–12.

[34] W. Feng *et al.*, “Joint 3D trajectory and power optimization for UAV-aided mmWave MIMO-NOMA networks,” *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2346–2358, Apr. 2021.

[35] Y. Sun, D. Xu, D. W. K. Ng, L. Dai, and R. Schober, “Optimal 3D-trajectory design and resource allocation for solar-powered UAV communication systems,” *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4281–4298, Jun. 2019.

[36] S. Hu, W. Ni, X. Wang, A. Jamalipour and D. Ta, “Joint optimization of trajectory, propulsion, and thrust powers for covert UAV-on-UAV video tracking and surveillance,” *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 1959–1972, 2021, doi: 10.1109/TIFS.2020.3047758.

[37] Stanford Artificial Intelligence Laboratory *et al.*, *Robotic Operating System*. [Online]. Available: https://www.ros.org

[38] DJI Co., *DJI Software Development Kits*. [Online]. Available: https://developer.dji.com/

[39] Docker Inc., *Docker*. [Online]. Available: https://www.docker.com/

[40] Google LLC *et al.*, *Kubernetes*. [Online]. Available: https://kubernetes.io

[41] *Status of Project IEEE 802.11ax*, IEEE 802 LAN/MAN Standards Committee, High Efficiency Wireless LAN Task Group. [Online]. Available: https://www.ieee802.org/11/Reports/tgax_update.htm

[42] IEEE 802.11ad, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band*. IEEE Standard, Dec. 2012

[43] *Status of Project IEEE 802.11ay*, IEEE 802 LAN/MAN Standards Committee, High Efficiency Wireless LAN Task Group. [Online]. Available: https://www.ieee802.org/11/Reports/tgay_update.html

[44] C. Fan, B. Li, C. Zhao, W. Guo, and Y. C. Liang, “Learning-based spectrum sharing and spatial reuse in mm-wave ultradense networks,” *IEEE Trans. Veh. Technol.*, vol. 67, no. 6, pp. 4954–4968, Jun. 2018.

[45] P. Zhou, X. Fang, X. Wang, Y. Long, R. He, and X. Han, “Deep learning-based beam management and interference coordination in dense mmWave networks,” *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 592–603, Jan. 2019.

[46] A. Geitgey, *Face Recognition*. [Online]. Available: https://github.com/ageitgey/face_recognition/

[47] “iperf3: A UDP TCP, and SCTP network bandwidth measurement tool,” [Online]. Available: https://github.com/esnet/iperf

**TAO YU** (Member, IEEE) received the M.E. degree in signal and information processing from the Communication University of China, Beijing, China, in 2010, and the Dr.Eng. degree in electrical and electronic engineering in 2017 from Tokyo Institute of Technology, Tokyo, Japan, where he is currently working as a Postdoctoral Researcher. His current research interests include aerial communication, wireless sensor networks and wireless energy transmission. He is the member of IEICE.

**YOSHITAKA TAKAKU** received the B.E. and M.E. degrees in electrical and electronic engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 2018 and 2020, respectively. His research interest focuses on aerial communication.

**YOHEI KAIEDA** received the M.E. degree in natural science and technology from Kobe University, Kobe, Japan, in 2007. His current research interests include wireless sensor networks and aerial communication.

**KEI SAKAGUCHI** (Senior Member, IEEE) received the M.E. degree in information processing from Tokyo Institute Technology, Tokyo, Japan, in 1998, and the Ph.D. degree in electrical and electronics engineering from Tokyo Institute Technology, Tokyo, Japan, in 2006. He is currently working with the Tokyo Institute of Technology in Japan, as the Dean with Tokyo Tech Academy for Super Smart Society and as a Professor with the School of Engineering. At the same time, he is working for oRo Co., Ltd. in Japan as the outside Director. His current research interests include 5G cellular networks, millimeter-wave communications, wireless energy transmission, V2X for automated driving, and super smart society. He was the recipient of the the Outstanding Paper Awards from SDR Forum and IEICE, in 2004 and 2005, respectively, and three Best Paper Awards from IEICE communication society in 2012, 2013, and 2015. He was also the recipient of the Tutorial Paper Award from IEICE communication society in 2006. He is a fellow of IEICE.