Technology Reports (Special Articles)

5G | Edge AI | Self-driving

Special Articles on AI—Expansion of AI Technologies to Diverse Industries and Basic Technologies Supporting AI Applications—

Self-driving Support System toward Personal Mobility Using Edge AI-compatible 5G Device

Communication Device Development Department Makoto Takahashi Yuuki Nakazawa Tomoya Moribe Kenya Ikeda

Expectations are growing for diverse self-driving services using personal mobility vehicles and service robots with the hope of solving a variety of social problems brought on by an aging society and labor shortages. Ensuring safety will be of prime importance in implementing these services in society—obstacle-avoidance performance needs to be improved and remote-steering services need to be supported. In response to these needs, NTT DOCOMO has developed technologies for obstacle detection on an edge device with high real-time performance and for secure and low-latency remote steering using 5G communications in a closed environment. These technologies support safe self-driving thereby contributing to the implementation of self-driving services in society.

1. Introduction

The total population of Japan as of October 1, 2020 was 125,710,000. As part of this figure, the population of those 65 years of age or over came to 36,190,000 bringing the percentage of elderly people in Japan to 28.8%. Statistical data from the Cabinet Office, Government of Japan indicates that approximately one in 2.6 people will be 65 years of age or over by 2065 [1]. In response to the aging society, there are expectations in Japan for the social implementation*1 of labor-saving services such as self-driving wheelchairs enabling the elderly to move about freely and self-driving robots for making last-mile deliveries to one's home. Ensuring safety in the social implementation of self-driving services is of prime importance, so initiatives for improving safety are needed. In addition to distance sensors such as conventional Light Detection And Ranging (LiDAR)*2, these will include an obstacle-detection function capable of instantaneous response using image recognition technology*3 operating on an edge device and support for a remote-steering mode to enable self-driving in situations where autonomous driving*4 is difficult.

With a view to safe self-driving in various types of personal mobility*5 vehicles and service robots, this article describes a self-driving support system toward personal mobility achieved by using a remote-steering system constructed in NTT DOCOMO's Multi-access Edge Computing (MEC)*6 environment in combination with an Edge AI*7-compatible 5G device.

2. System Overview

The overall configuration of a self-driving support system developed by NTT DOCOMO toward personal mobility is shown in **Figure 1**. This system consists of three functional blocks: (1) "Edge AI-compatible 5G device" as an add-on to various types of personal mobility vehicles and service robots, (2) "Edge AI applications" for performing processing requiring real-time performance on that device, and (3) "remote steering system" that supports the monitoring and driving of personal-mobility self-driving vehicles. The system takes on a configuration that can be used not for the development of personal mobility itself such as wheelchairs but rather as an add-on to vehicles developed by partner companies. It enables self-driving support that enhances safety for diverse types of personal mobility vehicles and service robots.

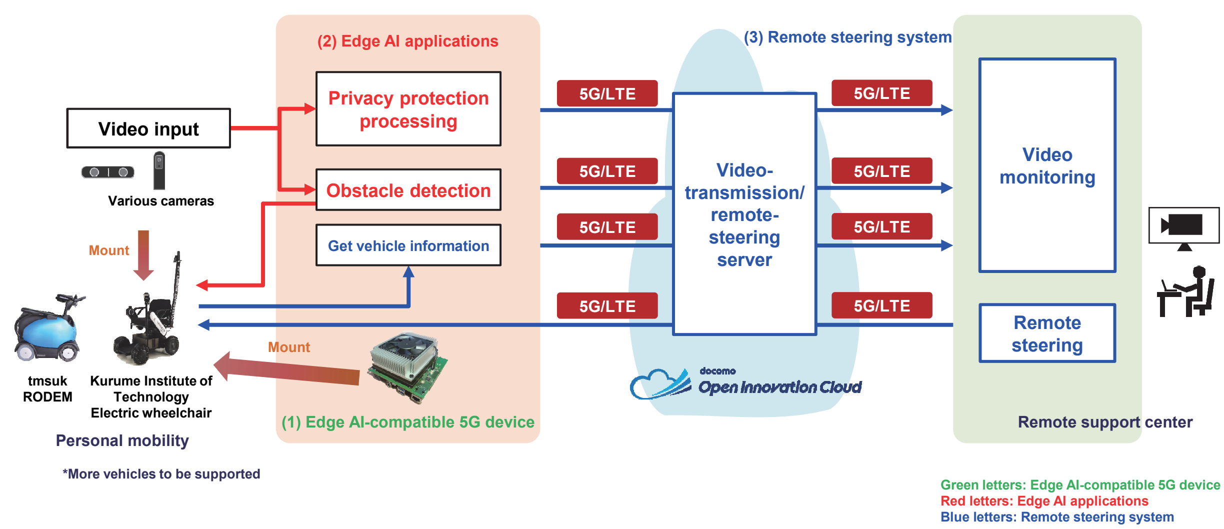

Figure 1 Overall configuration of a "self-driving support system toward personal mobility"

*1 Social implementation: In this article, the state in which a mechanism for solving social issues penetrates social life and continues to be used.

*2 LiDAR: Sensor technology that irradiates objects in the surrounding area with light such as near-infrared light and enables object detection by reflected light.

*3 Image recognition technology: Technology for mechanically understanding images and extracting meaning using image processing technology, machine learning technology, etc.

Specifically, this system will enable appropriate vehicle control such as by recognizing that a certain obstacle in the vicinity of a personal mobility vehicle is a person to be avoided. It can also enhance the safety of a self-driving system that uses only distance sensors such as LiDAR. Furthermore, with a view to the social implementation of self-driving services, there will be a need for a function that can remotely support self-driving to improve the safety of autonomous driving technology in the field of personal mobility and to deal with unforeseen problems during autonomous driving. This system transmits on-site video to a remote support center, which is equipped with a function for monitoring vehicle conditions. In this way, it becomes possible for a manager at the remote support center to perform remote steering in various situations, such as when the personal mobility vehicle loses self-position during autonomous driving due to a crowd of people or obstacles in the periphery or when the vehicle moves into a location in which no digital map exists during autonomous driving. This system will also consider the need for privacy in the case of people captured by on-site video, so it will enable the faces of any persons appearing in the video transmitted to the remote support center to be preprocessed by mosaic masking.

The cloud environment used for constructing this system is the DOCOMO Open Innovation Cloud® [2], an MEC environment provided by NTT DOCOMO. A direct connection to this cloud can be made using the 5th Generation mobile communications system (5G), which enables high-definition and low-latency video transmission unique to 5G as well as high security by not passing through the Internet.

3. Edge AI-Compatible 5G Device

3.1 Device Overview

This device consists of a high-performance processor, 5G module, and carrier board*8 equipped with various interfaces as described below. On applying this device to personal mobility, it must be capable of high-load processing as in the case of image recognition on an edge device and of large-capacity, low-latency communications for transmitting high-definition video. It must also be relatively easy to incorporate in the vehicle. By combining the constituent elements described below, this device can be used to support self-driving in personal mobility.

3.2 High-performance Processor Capable of Edge AI Processing

Autonomous driving in personal mobility requires multiple types of AI processing on an edge device excelling in real-time performance. While AI processing such as image recognition is generally executed on the cloud due to its superior machine performance, this device mounts the Jetson AGX Xavier [3] processor for edge computing*9 from NVIDIA enabling AI processing on an edge device. It also supports mounting of the Jetson Xavier NX [4] processor featuring lower power consumption and a lower price. The user can therefore choose which of these two processors fits the target application.

3.3 5G Module for High-speed, Large-capacity, Low-latency Communications

We adopted the Telit FN980 [5] module supporting 5G communications for which interoperability

---

*4 Autonomous driving: A function for driving to a destination while estimating self-location using a digital map and distance sensors, motion sensors, etc. mounted on a personal mobility (see *5) vehicle.

*5 Personal mobility: In this article, an electrically driven vehicle for personal use.

*6 MEC: A mechanism of installing servers or storages within a carrier network, at locations near users.

*7 Edge AI: In this article, AI processing not only in an MEC environment but also on actual devices.

*8 Carrier board: Hardware having essential input/output interfaces.

tests with the NTT DOCOMO network have been completed. In the area of self-driving support for personal mobility through remote steering, measures have been taken to improve safety by using multiple cameras to provide high-definition video, 360-degree video, etc., but the deployment of 5G has made the transmission of such video all the more feasible.

This module also supports an external antenna so that stable communications can be achieved through an antenna external to a personal mobility vehicle while embedding this device in the interior of the vehicle. The module also adopts a general-purpose M.2 connector*10 as a communication-module interface to ensure connectivity with various types of equipment.

3.4 Carrier Board Equipped with an Interface for External Equipment Connections

We developed a carrier board to integrate the high-performance processor and 5G module described above. This board incorporates abundant interfaces such as a serial communications interface and High-Definition Multimedia Interface (HDMI) input (Jetson AGX Xavier only) as listed in **Table 1**. The idea here is to enable connections to smartphones and diverse types of sensors including external cameras, which is difficult for ordinary routers. Moreover, envisioning the use of this device as an add-on to personal mobility vehicles, it has been given a form that emphasizes ease of embedding. This device has general-purpose features that support embedding in various types of vehicles and not just specific types of personal mobility.

4. Edge AI Applications

4.1 Overview of Applications

With the aim of improving the safety of self-driving

Table 1　Specifications of carrier board

Size	120 × 120 mm
Operating temperature	-20 – +80℃
Operating humidity	10 – 90%
Interfaces	・1×HDMI Type A (output) ・1×USB3.1 Type C ・2×USB3.1 Type A ・1×USB2.0 Micro B (OTG) ・1×RJ-45 for GbE ・1×Micro SD slot ・1×DC-in 9 – 19V (6 pin Euroblock) ・1×M.2 B-Key 3052 (Telit FN980m 5G/LTE) ・1×M.2 M-Key 2280 (SC710N1 M2 HDMI) (1×HDMI Type A input) ・1×M.2 M-Key 2280 (enables addition of M.2 external storage) ・1×nano SIM slot
Extension pin headers	・1×CANBUS ・1×UART 3.3V/5V TTL (JST-GH 6 pin) ・1×UART 3.3V/5V TTL (DF-13-6 pin) ・1×RS-232/1×I2C/5×GPIO ・1×MIPI CSI connector (120 pin) ・1×front panel (reset/recovery/power ON)

*9 Edge computing: Technology that distributes edge servers closer to the users to improve response and reduce latency.

*10 M.2 connector: A connection terminal applicable to thin-type, high-performance devices.

in personal mobility, NTT DOCOMO has developed applications to be run on an edge device. These include peripheral video recognition to avoid collisions with obstacles, which requires real-time performance, and privacy protection processing with respect to persons caught in video transmitted for remote-steering purposes.

4.2 Detection of Obstacles in Vicinity of Personal Mobility Vehicle

In autonomous driving, there is a need to accurately determine the presence of any obstacles in the vicinity of the personal mobility vehicle. To this end, the system performs image recognition and analysis using stereo-camera video to identify the type of obstacle and calculate the distance to it. Furthermore, if the obstacle happens to be a moving object such as a human being, the system predicts the direction and amount of movement. The information output from these various forms of video analysis will be instantaneously passed from the edge device to the personal mobility vehicle without any cloud intervention thereby enabling fast vehicle control and safer self-driving.

To improve the operability of a remote operator and to improve system safety in remote steering, the system performs 360-degree image recognition and analysis about the personal mobility vehicle and displays the results on a screen at a remote site. In this way, the remote operator can visually ascertain whether someone is in the vicinity, and if so, at what position.

4.3 Protection of Private Information

When transmitting video from a medical facility, for example, there are cases in which privacy protection is required with respect to any persons that happen to be included in the transmitted video. To deal with such a situation in this system, we made it possible to perform mask processing on the edge device itself with respect to any persons detected in the video before transmitting that video to the remote operator. Here, considering that total masking of a person may impair the remote operator's visibility, the system performs mask processing on only the face of a detected person. However, when a person's face itself becomes the target of image recognition and processing, there is a tendency for the accuracy of privacy protection to drop since the target area is relatively small. For this reason, the system uses person detection and face detection in a stepwise manner.

4.4 Reducing Latency in Transmitting Video and Control Information

In autonomous driving and remote steering of a personal mobility vehicle in which collisions must be prevented and safety ensured, it is vitally important to achieve low latency in addition to dealing with high-load processing. Additionally, given that multiple AI processes are being performed in parallel on an edge device, the limited computational resources on that device must be efficiently used. Consequently, to achieve low-cost and speedy development of this system, we used Yolo [6], an open-source AI framework*11, and devised measures to accelerate the processing of Edge AI applications.

In inference processing, this system forgoes single-precision floating-point operations*12 and adopts half-precision floating-point operations*13 instead. For example, in the case of a recognition target

*11 **Open-source AI framework:** A compilation of diverse programs to simplify AI development work is called an AI framework. At present, a variety of AI frameworks can be used without charge as open source software.

*12 **Single-precision floating-point operations:** 32-bit floating-point operations abbreviated as FP 32.

*13 **Half-precision floating-point operations:** 16-bit floating-point operations abbreviated as FP 16.

having a size comparable to that of a human being, it is known that the use of single-precision floating-point operations, though improving computational accuracy, does not greatly improve inference accuracy. We therefore decided to use half-precision floating-point operations since they can perform inference processing faster than single-precision floating-point operations. We also optimized each processing parameter in Edge AI applications, optimized the buffer size for decoding, and performed downsampling*14 of input images. In these ways, we made more efficient use of all computational resources running on the edge device and achieved low-latency operations.

5. 5G Remote Steering System

5.1 System Overview

This system consists of an Edge AI-compatible 5G device (Fig. 1 (1)), video-transmission/remote-steering server constructed on the DOCOMO Open Innovation Cloud, web/iPad*15 application for monitoring, and a controller for steering (Fig. 1 (3)). In video transmission, the edge device transmits the video and various types of sensor information created by an Edge AI application to the remote support center via the server for display on the web/iPad application for monitoring purposes. In remote steering, the steering information input by the controller is transmitted via the server to the edge device, which sends it on to the personal mobility vehicle. Here, communications are performed via the 5G/LTE network using a direct connection to the server on the DOCOMO Open Innovation Cloud. Issues surrounding the practical implementation of remote steering as uncovered during the study of this system are listed in **Table 2** and discussed below.

5.2 Low Latency and High Security Using the DOCOMO Open Innovation Cloud

From the viewpoint of low latency, operations from the input of video to the input/reflection of steering information must be completed within 500 ms so that the operator on the remote-support-center side can perform remote steering safely. We satisfied this requirement for low latency within 500 ms by devising measures for processing on the edge device as described above as well as by combining the DOCOMO Open Innovation Cloud, NTT DOCOMO's original MEC platform, with a direct cloud connection over a dedicated line. Constructing this remote steering system on the DOCOMO Open Innovation Cloud in this way makes for improved low-latency characteristics compared with the use of a public cloud*16 on the

Table 2 Issues in achieving a practical remote steering system

(1) Low latency	Low latency from checking conditions to reflecting steering is essential.
(2) 360-degree safety check	Visual recognition of people or objects approaching from the side or rear is essential.
(3) Vehicle information check	Monitoring of vehicle information such as remaining battery capacity is essential.
(4) Flexible controller support tailored to vehicle	Supporting a specific controller tailored to the vehicle is essential to enhancing system extensibility.

*14 Downsampling: The process of lowering the resolution of an image to reduce computational complexity.

*15 iPad: Apple, the Apple logo and iPad are trademarks of Apple Inc. registered in the United States and other countries. TM and ©2020 Apple Inc. All rights reserved.

*16 Public cloud: Cloud computing services that anyone can use over the Internet.

Internet. In trials conducted with vendors, we found that the time taken from transmitting video by an Edge AI application including privacy protection processing to its display at a remote site is about 300 ms while the time taken to transmit steering information is about 100 ms for a total of about 400 ms, which confirms that safe remote steering could be achieved within 500 ms. Another advantage of this system in addition to low latency is that the communication path has a closed configuration not open to the NTT DOCOMO network. The system is therefore separated from the outside, which reduces the risk of unauthorized access thereby achieving high-security characteristics.

5.3 Visualization of Video and Sensor Information

Next, from the viewpoint of ensuring safety during remote steering, it is important to not only secure a field of view in the forward direction but to also visually recognize persons and objects approaching from the side or the rear. Additionally, information on the condition and orientation of the personal mobility vehicle obtained from battery information, sensors, etc. must be monitored to ensure safe driving. These problems are addressed by displaying both webcam video in the forward direction and omnidirectional video from a 360-degree camera on the web/iPad application for monitoring. There is also a function for sending arbitrary text and numerical information from the edge device to the server for visualization on a web application (**Figure 2**). As for the 360-degree camera video, the viewpoint can be controlled by a drag operation so that the remote operator can zoom in on a particular section of the surrounding environment. Also provided is a radar-display function for presenting positional information on people detected by an AI application from 360-degree camera video. Additionally, with regard to the above

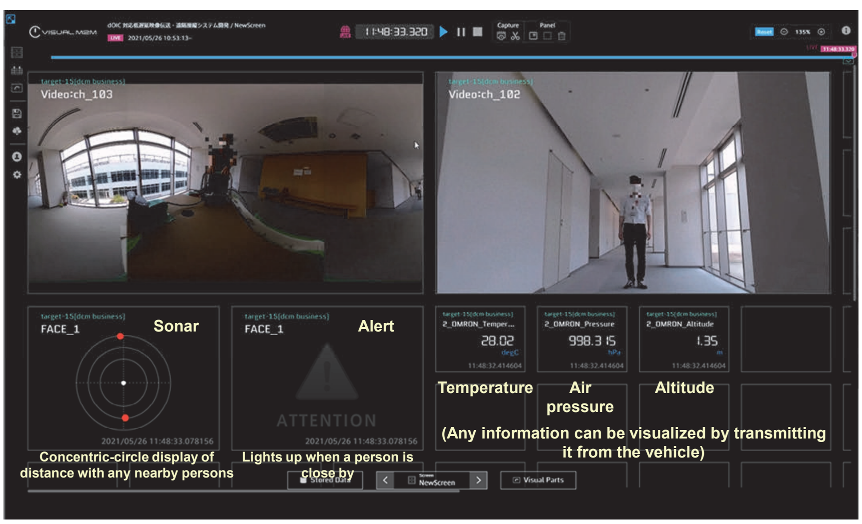

Figure 2　Screenshot of video-transmission/remote steering system

function for visualizing arbitrary data, the control application of each personal mobility vendor may transmit any information targeted for visualization by using a specific Application Programming Interface (API)*17. Each vendor may freely select the information to be visualized, the types of sensors to be used, etc.

5.4 Flexible Support of Multiple Mobility Controllers

Next, from the viewpoint of system operability, it must be possible to use diverse types of controllers specific to various types of personal mobility vehicles. In response to this problem, the specifications of this system allow for transparent transmission of steering information to the server instead of having to convert that information to some type of unique format. Here, the system transmits steering information with low latency via the DOCOMO Open Innovation Cloud so that remote steering can be achieved by incorporating that steering information in the control application developed by each personal mobility vendor. The vendor only needs to select the controller equipment conforming to the control application, which enables the system to provide flexible support for controllers applicable to each personal mobility vehicle.

6. Conclusion

With the aim of solving social problems brought on by an aging society and labor shortages, this article described a self-driving support system toward personal mobility that will contribute to safe self-driving of personal mobility vehicles and service robots.

It described, in particular, the features of an "Edge AI-compatible 5G device" that connects to a webcam and 360-degree camera and achieves image recognition processing and low-latency transmission of high-definition video for remote monitoring through 5G communications, "Edge AI applications" that enable obstacle detection for real-time vehicle control and privacy-protection processing against transmitted video to be performed on an edge device, and a "remote steering system" using the DOCOMO Open Innovation Cloud that enables secure and low-latency remote monitoring and steering.

A variety of trials have been held in this area and some commercial services on a chargeable basis have begun to be provided. However, the cost incurred by a total system including hardware and software is still high, so to make further inroads in society, the service-provision price must be reduced and more improvements must be made in system safety. Going forward, it will become increasingly important to reduce costs by developing a general-purpose system that incorporates low-cost and commercially competitive and available equipment and technologies in addition to NTT DOCOMO's 5G network and MEC platform. NTT DOCOMO will promote the creation of new value and the social implementation of new technologies in collaboration with co-creation partners toward the solution of diverse social problems.

REFERENCES

[1] Cabinet Office, Government of Japan: "2021 White Paper on the Aging Society (Summary), Section 1 Situation of the Aging Population," (In Japanese).
https://www8.cao.go.jp/kourei/whitepaper/w-2021/

*17 API: An interface that enables functions provided by an OS or middleware to be used by other application software.

gaiyou/pdf/1s1s.pdf
[2] NTT DOCOMO: DOCOMO Open Innovation Cloud (In Japanese).
https://www.nttdocomo.co.jp/biz/service/doic/
[3] NVIDIA: "Jetson AGX Xavier."
https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-agx-xavier/
[4] NVIDIA: "Jetson AGX Xavier NX."
https://www.nvidia.com/en-us/autonomous-machines/embedded-systems/jetson-xavier-nx/
[5] Telit: "FN980."
https://y1cj3stn5fbwhv73k0ipk1eg-wpengine.netdna-ssl.com/wp-content/uploads/2021/02/Telit_FN980-FN980m_Datasheet.pdf
[6] J. Redmon, S. Divvala, R. Girshick and A. Farhadi: "You Only Look Once: Unified, Real-Time Object Detection," Cornell University, May 2016.
https://arxiv.org/abs/1506.02640