NTT DOCOMO

Technical Journal

Vol.22 No.4 | Apr. 2021

DOCOMO Today

- R&D for Continuous Creation of New Business Value

Technology Reports (Special Articles)

Special Articles on AI for Improving Services and Solving Social Problems

- GOLFAI: Golf Swing Analysis Service Using Image Recognition Technology
- Technologies for Achieving Entertaining Dialogue and Friendly Chatbots
- AI-based Optimization of Bike Relocation in Docomo Bike Share Service
- Smartphone Log-based Stress and Cognitive Performance Estimation Technology

Technology Reports

- AR/MR Cloud Technology to Provide Shared AR/MR Experiences across Multiple Devices
- History of “Densoku-cars” for Research and Development Supporting the Evolution of Mobile Communications—Unique Vehicles Used in 1G – 5G Mobile Radio Field Experiments—

R&D for Continuous Creation of New Business Value

Executive Vice President
R&D Innovation Division
Naoki Tani

In Japan, fifth-generation mobile communications system (5G) services were launched in March 2020. Through co-creation with business partners, NTT DOCOMO has been creating 5G solutions that can contribute to workstyle reform, advanced industrial systems, and attractive urban development. From here on, we can expect 5G to evolve even further and to support an even greater number of use cases. In retrospect, a new generation of the radio system for mobile communications has appeared roughly every ten years and a new wave of services using these systems has occurred about every 20 years. The car phone launched in December 1979 became increasingly smaller over a 20-year period eventually morphing into a mobile phone for voice communication that spread throughout society. Next, the launch of the i-mode service in February 1999 led to sophisticated terminals such as i-mode mobile phones and subsequent smartphones and laid the foundation for expanded mobile multimedia that could provide an even greater diversity of information at your fingertips. Going forward, then, what kind of world can we expect from 5G following its start with pre-commercial services in September 2019? According to reference materials of the Radio Policy 2020 Roundtable [1] held by the Ministry of Internal Affairs and Communications (MIC), 5G is expected to contribute to all sorts of industries and to create new business value with an economic impact estimated to be about 47 trillion yen in Japan alone.

What issues must we face to achieve this world? I believe that we can sum these up with the keywords of "social change" × "technology evolution" × "ideas and design." First, in the area of social change, we must contribute to economic development while solving key social issues such as dealing with Japan's super-aging society and creating a non-contact remote society as a countermeasure to COVID-19 infections. Next, technology evolution refers to the remarkable evolution of Information and Communications Technology (ICT) in the form of 5G networks, Internet of Things (IoT), virtual space, Artificial Intelligence (AI), cloud computing, and edge computing and to the development of advanced devices such as robots, drones, and self-driving cars. It also refers to the growing impact of digital platformers typified by GAFA*1. Here, support for a "digital transformation," that is, the use of technology evolution to solve the above issues surrounding social change, should accelerate corporate reorganization and cross-industry collaboration, promote the appearance of new digital players, and drive workstyle reform. All in all, these changes should lead to solutions to social issues and the creation of new added value. However, according to the World Digital Competitiveness Ranking 2020 [2] put out by the Switzerland-based International Institute for Management Development (IMD), Japan received a ranking of 27th in overall performance among 63 countries around the world. While strong in areas such as mobile broadband usage and investment in new technologies, low digital/technological skills, and poor business agility and delay in data usage in the business domain were cited as weak points.

Finally, in light of the above, "ideas and design" are particularly important. At NTT DOCOMO's R&D Innovation Division, we are engaged in R&D activities toward the creation of a "well-being society" as a vision of the future. This will be a society that provides an ever-wider range of possibilities so that each and every person can reach their full potential in a mutually supporting community. To achieve this, we will promote value creation within a "cyber-physical fusion" framework while promoting the simultaneous evolution of network technologies and service technologies. Cyber-physical fusion begins by converting humans, things, and events in physical space into information. It then acquires and accumulates data in cyber space via the network based on that information and uses that data to forecast the future, discover knowledge, and actuate the physical space. Repeating these activities in a loop leads to value creation beyond the traditional boundaries of the corporate and industrial world. At present, though, while many mechanisms for accumulating and using digital data, video, etc. in cyber space are appearing throughout the world, there is still an insufficient amount of knowledge for connecting those mechanisms to fundamental business reform, creation of added value, etc. in physical space. In this regard, I believe that there are still many areas that need to be studied.

We are committed to achieving cyber-physical fusion and a new era of value creation through our R&D efforts while promoting the evolution of 5G as a source of great possibilities.

REFERENCES

[1] Ministry of Internal Affairs and Communications (MIC): "Radio Policy 2020 Roundtable Reference Materials," (In Japanese). https://www.soumu.go.jp/main_content/000420824.pdf

[2] IMD: "The IMD World Digital Competitiveness Ranking 2020 Results." https://www.imd.org/wcc/world-competitiveness-centerrankings/

*1 GAFA: Acronym for Google, Apple, Facebook, and Amazon, the four leading American IT companies.

NTT DOCOMO
Technical Journal
Vol.22 No.4 Apr. 2021

[Contents]

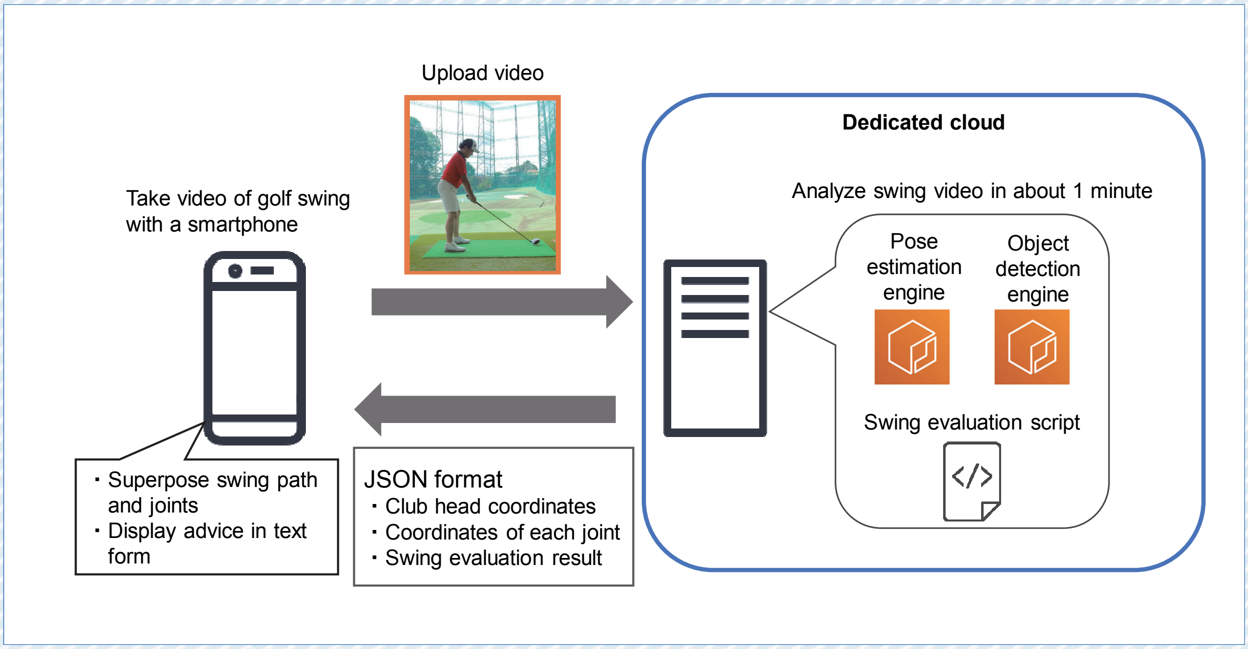
(P.14)

(P.23)

Technology Reports

(P.43)

(P.49)



Technology Reports (Special Articles) GOLFAI: Golf Swing Analysis Service Using Image Recognition Technology (P.4)
System configuration

Technology Reports (Special Articles) Image Recognition Deep Learning AI

Special Articles on AI for Improving Services and Solving Social Problems

GOLFAI: Golf Swing Analysis Service Using Image Recognition Technology

Service Innovation Department Katsuaki Kawashima Toshiki Sakai
Communication Device Development Department Muhd Hilmi Bin Shapien Yuki Ito
Business Creation Department Mizuki Watabe

Golf lesson apps that can automatically analyze one's golf swing are becoming increasingly popular. Many of these apps, however, require high-priced dedicated sensors that can deter their use for beginners and intermediate players. To provide low-cost and simplified golf swing analysis, NTT DOCOMO developed GOLFAI, a service that can analyze golf swings using only the video taken with a smartphone. With GOLFAI, a golf swing can be analyzed with a single smartphone and the user can receive personalized advice on one's swing.

In addition, we speeded up the development and provision of this app by adopting an in-house UI/UX development process.

1. Introduction

GOLFAI [1] is an NTT DOCOMO service that automatically analyzes golf-swing video uploaded by the user through image recognition technology*1 and uses the results of that analysis to provide free advice to the user. It has been available from the App Store since March 2020 as a service targeting mainly beginners and intermediate players who would like to improve their golf swing.

A golfer can attend golf school and receive personalized instruction directly from a golf instructor to improve one's level of play, but lesson fees, travel time to and from the school, etc. can place a burden on the student. To ease this burden, recent years have seen an increasing number of golf lesson apps such as Smart Golf Lesson [2] that can automatically analyze the user's swing. Many of these apps, however, require that high-priced dedicated sensors be attached to the golf club, which has presented a high hurdle to their use for beginners and intermediate players. Against this background, GOLFAI adopts image recognition using deep learning*2 to sense the body's joint positions and swing path from only video images captured by a smartphone as an alternative to dedicated sensors. In this way, GOLFAI enables low-cost and simplified golf-swing analysis compared with existing apps.

To help a person become better at golf, it is not simply a matter of determining whether that person's swing is good or bad—the user of a golf lesson app needs to be presented with more detailed information such as what part of the swing is bad and what kind of practice should improve the swing. In deep learning, it is generally difficult to describe the grounds for reaching a certain inference, and as a result, it is frequently difficult to obtain the information needed by the user simply on the basis of deep learning. GOLFAI, although using deep learning for sensing purposes, solves this problem by using a classic rule-based technique that incorporates the knowledge in expert-taught golf lessons for evaluating a swing.

In this article, we present specific examples of GOLFAI image recognition functions and describe our in-house UI/UX*3 development process that we adopted to speed up the delivery of this service to users.

2. GOLFAI System Overview

The GOLFAI system configuration is shown in **Figure 1**. In GOLFAI, the user begins by taking

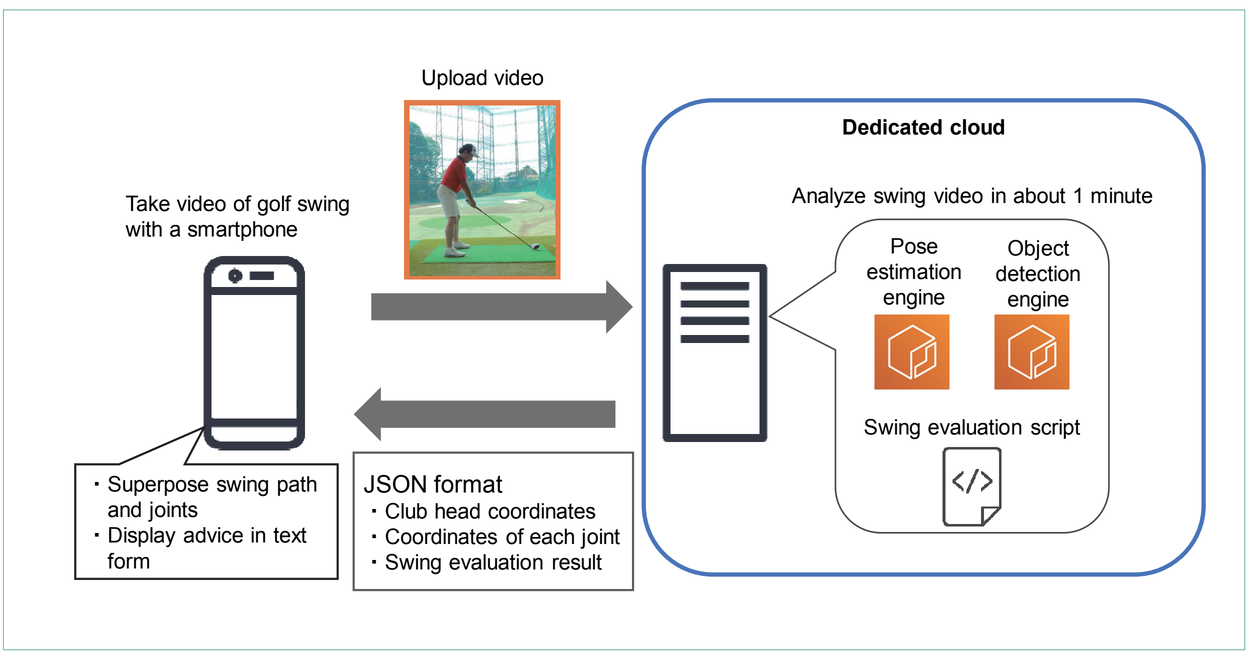


Figure 1　System configuration

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

*2　Deep learning: A type of machine learning that uses a multi-layered neural network (see *14).

*3　UI/UX: Abbreviations of "user interface" and "user experience."

video of his or her golf swing using a smartphone and uploading the video to a dedicated cloud. The GOLFAI system then performs image recognition processing against the uploaded video on a server on that cloud and returns the results of image recognition to the app in JavaScript Object Notation (JSON)*4 format. Finally, the app superposes those results of image recognition on the video or provides golf instruction by displaying advice in text form.

3. Image Recognition Overview

GOLFAI uses object detection and pose estimation as two types of deep learning for extracting from images the feature points needed for golf instruction. In applying deep learning, we independently collected about 1,000 clips of golf-swing video taken under a variety of conditions such as type of golf club and shooting location. Each clip is approximately 20 seconds long and is shot with Full HD resolution at 120 frames per second (fps)*5. We succeeded in improving the accuracy of feature-point extraction by performing fine-tuning*6 with this data.

GOLFAI uses extracted feature points and a rule-based technique to detect a frame*7 that corresponds to an important position in a golf swing and to categorize swing type. It then converts this information into a form that is easy for the user to understand and offers advice.

3.1 Feature Point Detection

1) Object Detection

Object detection refers to technology that estimates the position and class of a predefined object within an image. In GOLFAI, the system detects the golf club head and calculates the path of the swing by performing object-detection processing against all frames (**Figure 2** (a)).

In object detection technology, learning-based techniques using deep learning have reached high levels of accuracy in recent years. These techniques


(a) Object detection

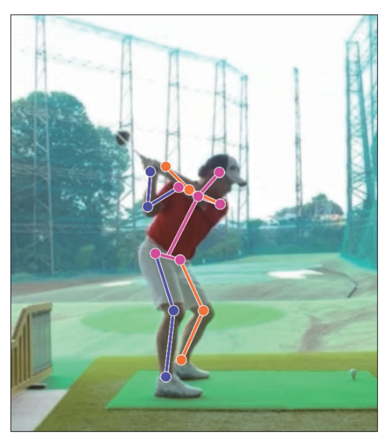

(b) Pose estimation

Figure 2 Example of extracting feature points

---

*4 JSON: A data description language based on JavaScript object notation.

*5 fps: Number of still images per unit time.

*6 Fine-tuning: A technique that sets parameters of a model that has already been learned once to initial values and then uses a different dataset to relearn that model and finely adjust those parameters.

*7 Frame: One of the many single still images that make up video.

can be broadly divided into the following two types.

(1) Two-stage detection framework

After preprocessing that proposes object candidate areas, this technique performs two-stage processing to estimate the class of object and area coordinates. This two-stage framework begins by generating candidate areas using a feature map*8 obtained through a selective search*9 or Convolutional Neural Network (CNN)*10. It then calculates features*11 from each candidate area and estimates object type and detailed coordinates using a support vector machine*12 or other type of classifier*13. In this framework, typical neural networks*14 include SPPNet [3], Fast RCNN [4], and Faster RCNN [5].

(2) One-stage detection framework

This technique estimates object candidate areas and object class and area coordinates using a single neural network. In short, it is a one-stage framework that directly estimates candidate-area coordinates and class in a single process. It has a simple processing structure that can perform all calculations including those for generating candidate areas using a single neural network. For this reason, a one-stage detection framework is said to be faster than a two-stage detection framework in terms of learning and inference. Here, typical neural networks include YOLO [6], SSD [7], and CornerNet [8].

For GOLFAI, we have adopted a one-stage framework as shown in **Figure 3** to reduce machine resources and processing time. This model has three detection layers each having the role of detecting an object of different size, that is, a large, medium, or small object. In this way, it becomes possible to perform high-accuracy object detection regardless of the size of the object in the image.

2) Pose Estimation

Pose estimation is technology for estimating the positions of human joints in an image in the form of two-dimensional or three-dimensional coordinates. The targets of estimation must be defined and learned beforehand, and in GOLFAI, the system calculates the two-dimensional coordinates of the joints needed for evaluating a golf swing such as elbows, wrists, and knees (Fig. 2 (b)). Pose estimation technology has been reaching high levels of accuracy owing to recent advances in learning-based estimation techniques using deep learning. In this regard, pose estimation technology using deep learning can be mainly divided into the following two approaches.

(1) Top-down approach

This is a technique that detects persons within an image using object detection technology and detects the joints of each of those persons. It's a simple technique that performs person detection and joint detection separately and that can improve the accuracy of pose estimation by improving the accuracy of person detection. Its weak point is that its computational cost increases in proportion to the number of persons in the image. Here, typical neural networks include DeepPose [9], Cascaded Pyramid Network [10], and High-Resolution Network [11].

(2) Bottom-up approach

This is a technique that first detects all human joints existing in the image and then

---

*8 Feature map: In this article, a multidimensional array obtained from the results of processing an input image by a CNN.

*9 Selective search: A technique that calculates object candidate areas by grouping similar pixels in an image.

*10 CNN: A type of neural network (see *14) that introduces processing for multiplying vectors of specific sizes while scanning a multidimensional array in the vertical and horizontal directions. CNN is widely used in image recognition.

*11 Feature: An amount (a numeric value) extracted from data to characterize that data.

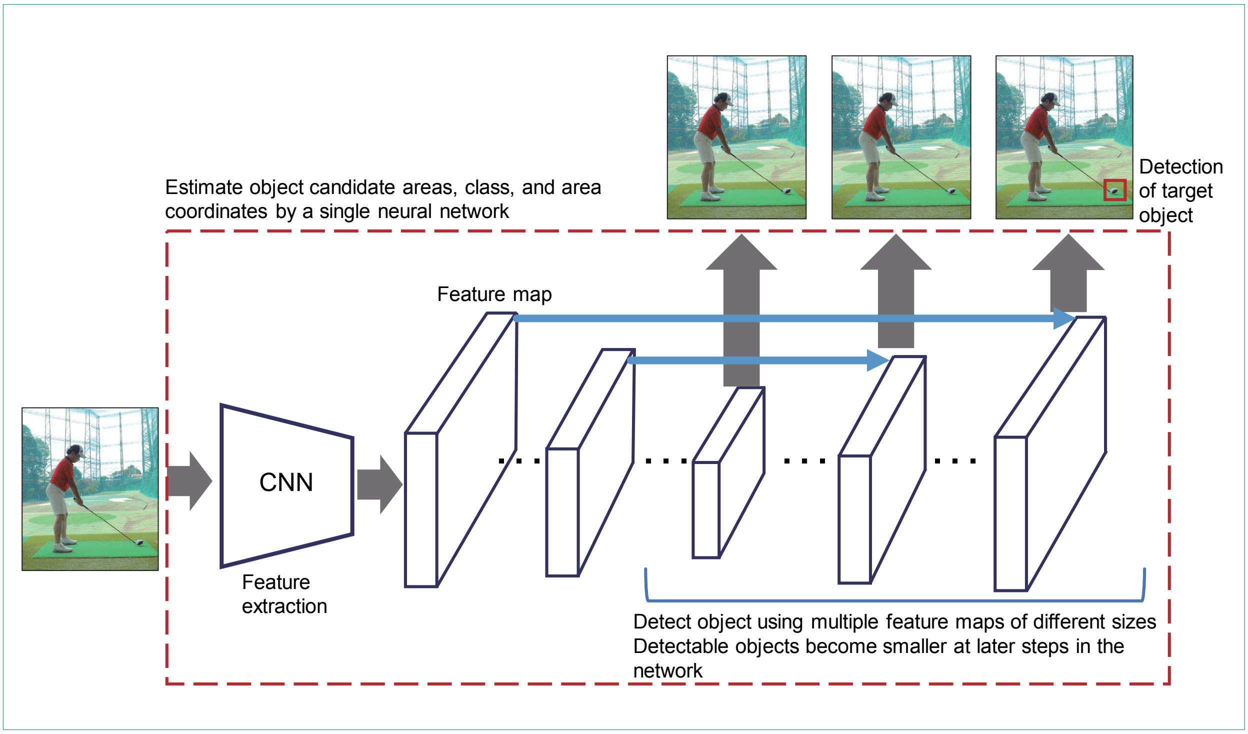


Figure 3 Overview of object detection processing

connects those joints on a person-by-person basis. Since it detects the joint points of all persons in the image at one time using a single neural network, inference speed is hardly changed by the number of persons present. On the other hand, accuracy may suffer even if joints have been correctly detected if failures occur in connecting the joints. Typical neural networks here include Deep Cut [12], Open Pose [13], and PersonLab [14].

For GOLFAI, we decided to use the bottom-up approach to prevent an increase in processing time and incorporated a model consisting of two neural networks to estimate joint positions and joint connections (**Figure 4**). In this model, learning not only joint position but joint orientation as well enables high-accuracy connection of detected joints for each person.

3.2 Swing Analysis

1) Estimation of Swing Position

GOLFAI detects the frames corresponding to the four positions of address (A), top (T), impact (I), and finish (F) considered to be key components of the golf swing. As shown in **Figure 5**, this enables the user to check one's swing by simply touching the A, T, I, and F buttons to move the playback position to the frame of the desired position. Although individual differences can be found in golf swings, GOLFAI focuses on standard elements of swing movement and estimates position in a rule-based manner.

*12 Support vector machine: A machine-learning method used in pattern recognition. It can be applied even to problems that are not linear separable by the "kernel trick" method.

*13 Classifier: An algorithm that classifies input into one of a number of predetermined classifications based on features.

*14 Neural network: A mathematical model that mimics the structure of the human brain.

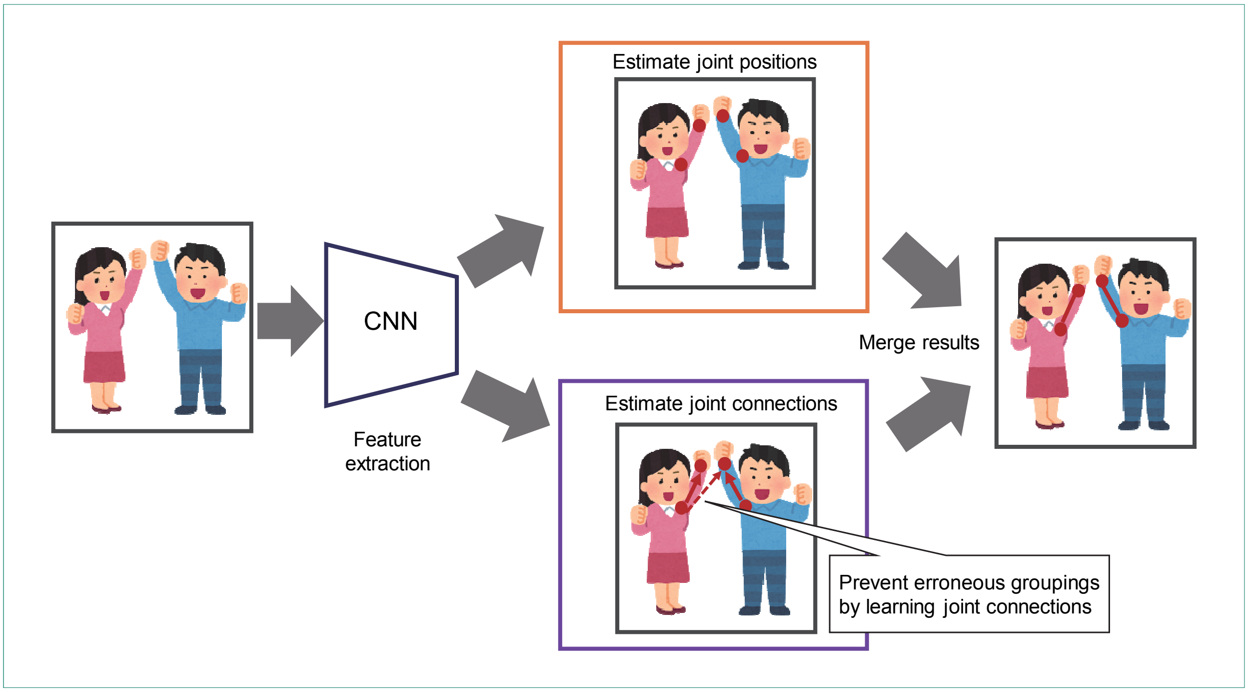


Figure 4　Overview of pose estimation processing

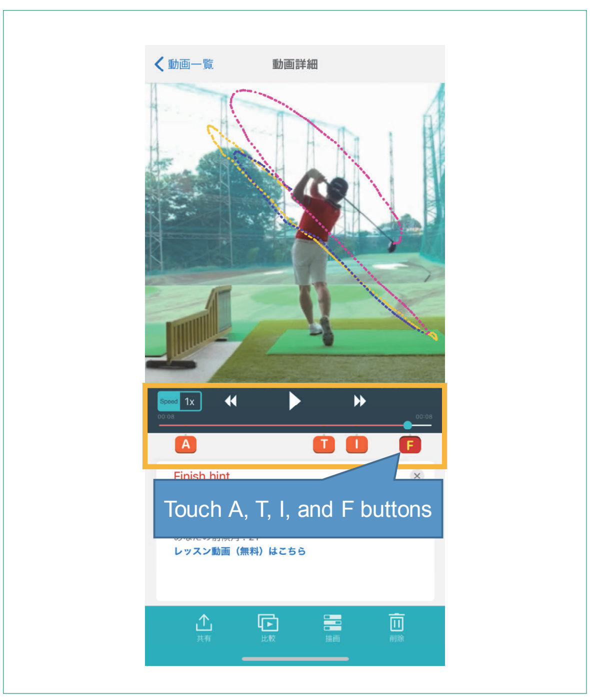


Figure 5　Example of position estimation results

2)　Swing Evaluation

This process evaluates a swing using the feature points calculated using object-detection and pose-estimation technologies. We developed a swing evaluation method by holding discussions with experts on swing evaluation points and their evaluation methods and by applying proprietary image recognition techniques. Through these discussions, we broke down expert knowledge into mathematical formulas and used a rule-based technique to evaluate swings. The following presents examples of evaluation points in GOLFAI.

(a) Maintaining body's forward tilt angle (**Figure 6** (a))

It is said that keeping the body's forward tilt angle fixed during a swing makes for a good swing. To evaluate whether a swing is good from this viewpoint, GOLFAI first

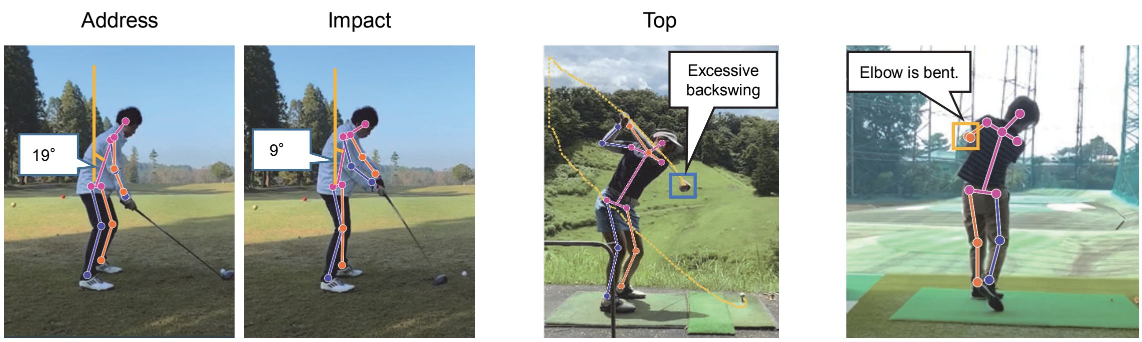


Figure 6　Example of swing evaluation

calculates the forward tilt angle for each position using the coordinates of the hip and the neck. Then, if the amount of change in the forward tilt angle between positions should exceed a fixed value, it offers the user advice on how to improve the swing. This fixed value is determined on the basis of expert knowledge.

(b) Overswing (Fig. 6 (b))

It is also said that an excessive backswing called an "overswing" is generally not good since it can induce instability in the swing path. To evaluate whether a swing is bad from this viewpoint, GOLFAI judges that an overswing occurs when the golf club head at top position is located on the ball side relative to the body at a point lower than the neck.

(c) Chicken wing (Fig. 6 (c))

A swing in which the left elbow (right elbow for a left-handed player) is bent in the follow-through is called a "chicken wing." According to experts, the elbow will be hidden by the user's body and out of view during the follow-through when observing elbow movement in a normal swing (no chicken wing) from a position lateral to the user. Based on this knowledge, GOLFAI judges that a chicken-wing swing occurs after impact when the left elbow (right elbow for a left-handed player) is detected in the image before the golf club head traverses the body.

In the above way, GOLFAI can offer the user advice in an easy-to-understand format by incorporating a rule-based technique in swing evaluation instead of using deep learning.

4. Shortening GOLFAI Development Period and Improving Ease-of-use by In-house UI/UX Studies

4.1　Service Development Issues

In a service app market that has become extremely competitive in recent years, how to provide an app with a superb UI/UX as quickly as possible to acquire users has become a major issue. We therefore felt the need to efficiently and quickly develop and market GOLFAI, a service that makes use of new technologies, while giving full consideration to the UI/UX.

However, service development at NTT DOCOMO currently takes five to six months on average from initial studies to service provision, so this time, we focused on studying UI/UX design and improving the output process (hereinafter referred to as "UI/UX process"). Improving the UI/UX process solves the following two issues.

- Conflicts in interpreting UI/UX

  To improve the UI/UX, a process of trial and error must be repeated to reach an agreement among stakeholders (on the app-development side) as to the value provided to the user while taking user response into account. It is difficult, however, to reach a common understanding in this way.

- Inefficiencies in communicating information

  Communication among stakeholders must be sufficiently considered, but when working with an outside design company, maintaining close communication and achieving a smooth development process is difficult.

As the number of stakeholders increases, exchanging information becomes more complicated and more time is taken up by communication activities, all of which can slow down the development process. Taking this into consideration, we tested an "in-house UI/UX process" in GOLFAI development with the aim of improving this process.

4.2 In-house UI/UX Process

In GOLFAI development, we studied and applied an in-house UI/UX process in the development department without consigning any work to an outside design company.

As shown in **Figure 7**, the development department created and proposed functions and a design based on concepts established at a workshop attended by all stakeholders. By assigning the UI/UX process to the development department, the number of stakeholders could be reduced and conditions at both the planning department and development vendor could be readily understood. This made it possible to enforce more effective and efficient proposals in a form that achieved a balance between "service requirements" and "development load." In this way, the development department took on the role of a mediator that could speed up proposals for specifications with the planning department and feasibility studies with the development vendor even when repeating the improvement cycle.

Specifically, the development department created a design proposal using Sketch*15 [15], a tool that is widely used for creating and delivering design data, and completed a final version of GOLFAI after repeating the improvement cycle multiple times. Here, at the stage with no operable app, screen transitions were simulated using Prott*16 [16], a tool that simplifies prototyping*17. This made it possible to check the design in a form close to actual operation from the initial development stage, which had the effect of speeding up the improvement cycle.

As a result of adopting an in-house UI/UX process in the development department, we were able to release GOLFAI in just four months. We

---

*15 **Sketch**: A vector graphics editor provided by Bohemian Coding in the Netherlands. It is generally used as a UI design tool for apps and the web.

*16 **Prott**: A prototyping tool specifically for app and web design provided by Goodpatch Inc. in Japan.

*17 **Prototyping**: An early sample, model, or release of a product to test and evaluate a concept.

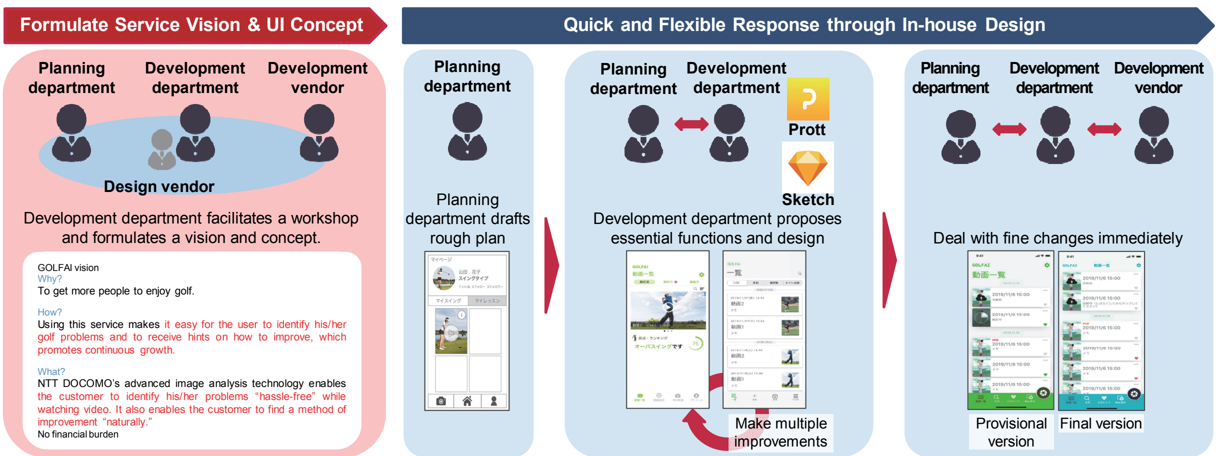


Figure 7　Development flow by an in-house UI/UX process

consider that the UI/UX study process contributed to this shortening of the development period by about one month. In short, an in-house UI/UX process can be highly effective if the objective is to release a new service as quickly as possible and if the project is small or medium in size in which the effect of reducing the number of stakeholders is high.

5. Conclusion

In this article, we described two key technologies used by GOLFAI: object detection technology for detecting the golf club head and pose estimation technology for estimating the positions of joints on the human body. We also described a swing evaluation method using the feature points extracted by those technologies and explained our adoption of an in-house UI/UX process for speeding up the delivery of this service to users. To help improve a golf swing, it is not sufficient to merely judge whether a swing is good or bad. It is also necessary to give specific advice such as what part of the swing is bad and what method can be used to improve the swing. This service provides the user with helpful advice by combining high-accuracy feature point detection using the latest deep learning techniques and expert knowledge. Going forward, we plan to develop image recognition technology that can achieve high-accuracy recognition even under poor shooting conditions as in a backlit scene. We also plan to study the application of this technology to other sports such as soccer and means of widening the application of this in-house UI/UX process.

REFERENCES

[1]　GOLFAI Official Site (In Japanese).

https://golfai.jp/

[2] SONY Smart Golf Lesson (In Japanese).
https://smartsports.sony.net/golf/product/1G/JP/ja/

[3] K. He, X. Zhang, S. Ren and J. Sun: "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," In ECCV, 2014.

[4] R. Girshick: "Fast R-CNN," In ICCV, 2015.

[5] S. Ren, K. He, R. Girshick and J. Sun: "Faster R-CNN: Towards Real Time Object Detection with Region Proposal Networks," In NIPS, 2015.

[6] J. Redmon, S. Divvala, R. Girshick and A. Farhadi: "You only look once: Unified, real time object detection," In CVPR, 2016.

[7] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Fu and A. Berg: "SSD: Single Shot Multibox Detector," In ECCV, 2016.

[8] H. Law and J. Deng: "CornerNet: Detecting Objects as Paired Keypoints," In ECCV, 2018.

[9] A. Toshev and C. Szegedy: "DeepPose: Human Pose Estimation via Deep Neural Networks," In CVPR, 2014.

[10] Y. Chen, Z. Wang, Y. Peng, Z. Zhang, G. Yu and J. Sun: "Cascaded Pyramid Network for Multi-Person Pose Estimation," In CVPR, 2018.

[11] K. Sun, B. Xiao, D. Liu and J. Wang: "Deep High-Resolution Representation Learning for Human Pose Estimation," In CVPR, 2019.

[12] L. Pishchulin, E. Insafutdinov, S. Tang, B. Andres, M. Andriluka, P. Gehler and B. Schiele: "Deepcut: Joint Subset Partition and Labeling for Multi Person Pose Estimation," In CVPR, 2016.

[13] Z. Cao, T. Simon, S.-E. Wei and Y. Sheikh: "Realtime Multi-person 2D Pose Estimation using Part Affinity Fields," In CVPR, 2017.

[14] G. Papandreou, T. Zhu, L.-C. Chen, S. Gidaris, J. Tompson and K. Murphy: "PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model," In ECCV, 2018.

[15] Sketch website.
https://www.sketch.com/

[16] Prott website.
https://prottapp.com

Technology Reports (Special Articles)
Chatbot Character Word Play
Special Articles on AI for Improving Services and Solving Social Problems

Technologies for Achieving Entertaining Dialogue and Friendly Chatbots

Service Innovation Department Yuiko Tsunomori Yutaro Shiramizu

Recent years have seen increasing demand for chatbots in all sorts of applications. In dialogue with chatbots whose purpose is essentially communication, it is important that the user finds that dialogue to be friendly and entertaining. To this end, we developed technology for reproducing specific characters and technology for engaging in word play such as *nazokake* riddles. These technologies make it possible to give chatbots an entertaining nature. We also developed a dialogue application running on a chat tool as a means of providing entertaining dialogue.

1. Introduction

Chatbots such as smart speakers and dialogue agent applications are becoming increasingly popular. Most of these accept various types of task requests such as "Please set an alarm" or "What will the weather be like today?" These chatbots are designed with this in mind, that is, to carry out a requested task, but in the case of "Shabette Concier," a voice agent released by NTT DOCOMO in 2012, many user utterances came to be input not with the aim of requesting a task but simply to communicate, i.e., to carry on a conversation. To respond to such user input, NTT DOCOMO developed a chat-oriented dialogue Application Programming Interface (API)*1 based on technology from NTT Media Intelligence Laboratories and released it for public use on the docomo Developers support (dDs) [1] website in 2013. Then, in 2018, NTT DOCOMO released the chat-oriented "katarai"

*1 API: An interface that enables software functions to be used by another program.

[2] service developed on the basis of this chat-oriented dialogue API in collaboration with Intermedia Planning, Inc. Rather than carrying out tasks, the purpose of this API and service is to communicate with the user, so it is important that such dialogue be friendly and entertaining to the user.

To achieve dialogue of this type, we developed character chat technology for reproducing specific characters and word play technology that can achieve chatbot dialogue having game-like characteristics as in posing riddles. These technologies make it possible to give chatbots an entertaining nature that makes them friendly and approachable to users. Additionally, to make it easy for users to enjoy dialogue having an entertaining nature (hereinafter referred to as “entertaining dialogue”), we developed an entertaining dialogue application that runs on a chat tool.

In this article, we describe the above character chat technology, word play technology, and entertaining dialogue application and present application examples.

2. Character Chat Technology

2.1 Overview

There is a high possibility that giving chatbots personalities can achieve dialogue that is friendly to users. For example, creating a character that a user likes or an original character can arouse a user’s interest.

Based on technology [3] from NTT Media Intelligence Laboratories, character chat technology reproduces the dialogue of arbitrary characters. This technology creates utterance-pair data (questions and answers) in the character style desired and constructs a dialogue system based on that data. In addition to constructing a new dialogue system, this technology can give personality to an existing dialogue system by combining the two systems.

Character chat technology achieves high-accuracy dialogue that reproduces a character at low cost with real-time characteristics. In use, it first extracts candidates by searching an utterance-pair database and then decides on a final utterance based on a score calculated by a variety of techniques. In this way, it can narrow down the candidates for an answer obtained from the initial search and achieve dialogue with a real-time feel even without a high-performance server.

2.2 Process Flow

The system configuration of character chat technology is shown in **Figure 1**. The system performs the following processing on input of a user utterance.

First, the utterance-pair search block searches for and extracts multiple answer candidates with respect to the input user utterance from an utterance-pair database that holds utterance pairs (question-answer pairs) for an arbitrary character. Specifically, this block searches the database for questions similar to the input user utterance and extracts a set of question-and-answer pairs as candidates corresponding to the user’s question.

Next, the utterance-score calculation block calculates multiple scores using a variety of techniques including a translation model and word2vec*2 model. In general, a translation model learns using a bilingual corpus*3 that assigns a correspondence between sentences in two languages (such as Japanese

*2 word2vec: A technique that analyzes text data and represents the meaning of each word in vector form.

*3 Corpus: A language resource consisting of a large volume of text and utterances, etc. collected and stored in a database.

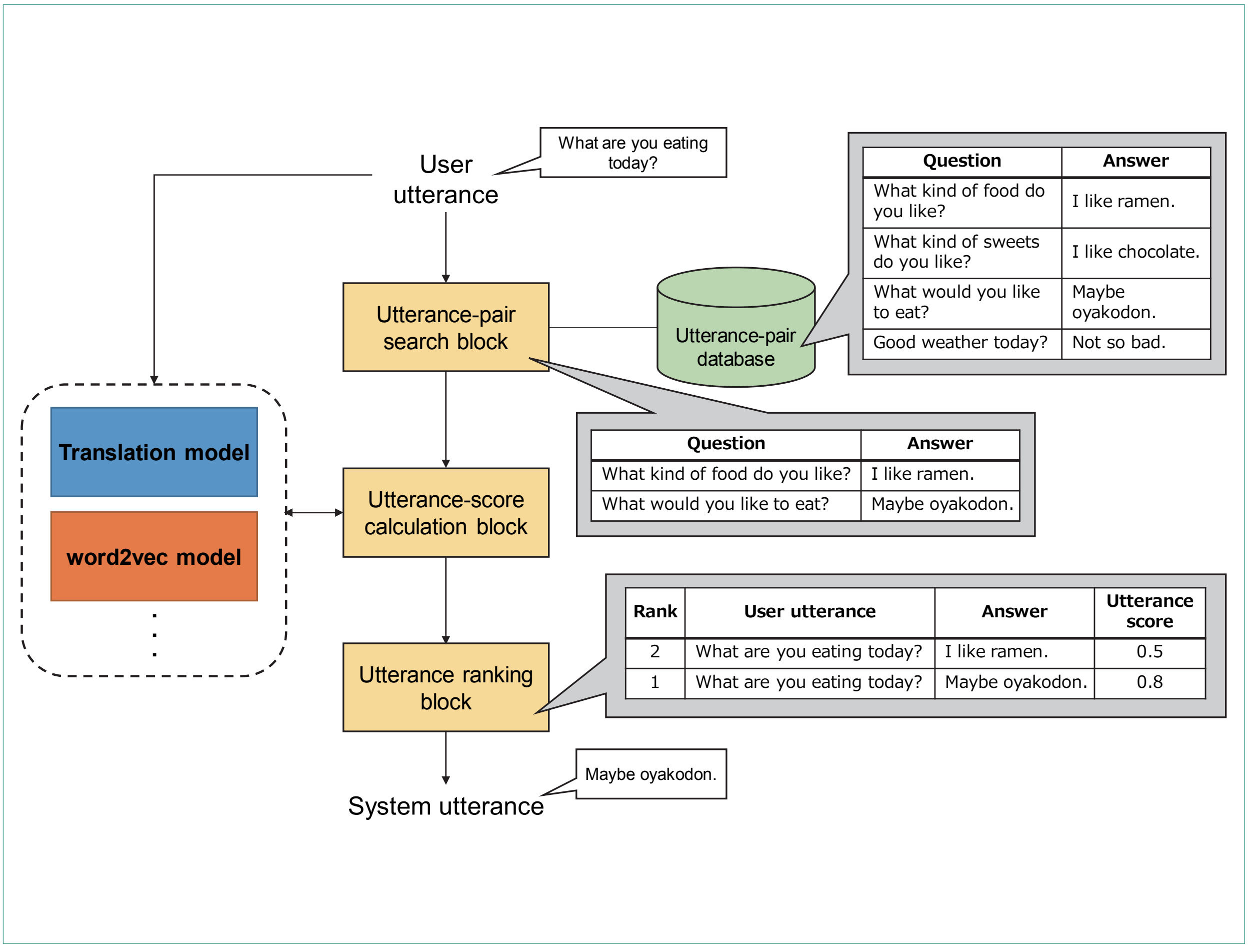


Figure 1 System configuration of character chat technology

and English). Our translation model, however, learns using a large volume of utterance pairs that assigns a correspondence between questions and answers. Specifically, for each question-answer pair consisting of the input user utterance and an answer candidate received as input, this translation model calculates a score on the naturalness of that user-utterance/answer-candidate pair. The word2vec model calculates a score on the similarity between the input user utterance and answer candidate.

Finally, the utterance ranking block integrates multiple scores and performs utterance ranking. The utterance with the highest score is output as the final system utterance (answer).

2.3 Application Example

"AI Jimmy" is the first commercial service to make use of character chat technology. Scheduled for provision by Laugh & Peace Mother Co. Ltd., AI Jimmy is an application that enables voice interaction with Artificial Intelligence (AI) that reproduces the Japanese entertainer Mr. Jimmy Onishi (**Figure 2**). That is, it can achieve dialogue that reproduces Mr. Jimmy Onishi by creating

Figure 2 AI Jimmy

and learning question-answer pairs by actually having Mr. Jimmy Onishi himself give answers to a variety of questions. For example, if the user inputs the question "How are you?" the application would respond in Mr. Jimmy Onishi's way of speaking by outputting "Me? I'm doing great—I can't stop!" Character chat technology can make answers sound all the more like Mr. Jimmy Onishi.

Voice interaction in AI Jimmy is achieved by using a DOCOMO AI Agent API and a speech synthesis model developed by NTT Media Intelligence Laboratories in addition to character chat technology. This DOCOMO AI Agent API features speech-recognition and speech-synthesis functions that enable voice interaction with the user. Here, speech synthesis in the DOCOMO AI Agent API makes use of an original speech-synthesis model developed in collaboration with NTT Media Intelligence Laboratories and trained with the speech of Mr. Jimmy Onishi.

3. Word Play Technology

3.1 Overview

Giving chatbots dialogue functions with game-like characteristics has a high possibility of heightening user interest and giving the user a feeling of "I'd like to talk more." This technology is being used to provide the four games of Yamanote Line, *Hyakunin Isshu* (traditional Japanese card game), *nazokake* riddles, and quiz.

1) Yamanote Line

The Yamanote train line is the famous loop line of Tokyo. Here, Yamanote Line is an interactive game based on specific themes such as "names of stations of the Yamanote line" in which the user gives answers such as "Shinjuku" and "Shibuya." You lose by duplicating an answer or giving an incorrect answer.

One issue that arises here when the user inputs an answer is "orthographical variants" (different notational forms). For example, if the user inputs the Japanese characters "新宿三丁目駅" corresponding to "Shinjuku-sanchome Station" in English (where the Chinese character 駅 = station), we can envision other forms of input with the same meaning such as "新宿三丁目" where the suffix駅/station is omitted or "新宿3丁目駅" where the Chinese numeral 三 meaning "three" is replaced by the Arabic numeral "3." We deal with this problem by automatically collecting information on orthographical variants from Wikipedia and regularly updating our dictionary. Additionally, as one means of enhancing game-like characteristics in dialogue, we establish an order for system answers according to the degree to which a name is known (name recognition) as calculated from the number of Wikipedia page views, link frequency, etc. For example, if the theme is "names of stations of the Yamanote line" as described above, the system would first answer with station names having relatively high name recognition such as "Shinjuku" and "Shibuya" and answer later with station names having relatively low name recognition such as "Tamachi."

2) *Hyakunin Isshu*

Hyakunin Isshu is a traditional Japanese card game based on a classical collection of one hundred short poems (each five lines in length) by one hundred poets called *Ogura Hyakunin Isshu*. In the basic format of this game, someone reads out the first three lines of one of those poems and the player looks for and picks up the card corresponding to the last two lines of that poem from among a set of cards each written with the last two lines of a different poem.

This word play function operates in a similar way. The system presents the user with a "reading card" showing the first half of a poem and multiple "pickup" cards showing the last half of different poems including the one corresponding to the first half of the poem shown on the reading card. The user selects one of those pickup cards presented by the system, which then informs the user whether the card selected showing the last half of a poem corresponds to the reading card showing the first half of a poem, i.e., whether the two cards correspond to the same poem (right/wrong answer).

In the "my daiz" member[*4] "Word Play" described later, the system displays a reading card while reading out the lines on that card by speech synthesis. Here, however, instead of directly uttering the text displayed on the reading card (written in historical kana orthography), a developer first makes some adjustments to that character string (such as replacing words with others having the same reading, inserting commas, etc. to enhance or diminish intonation) to achieve a natural card reading appropriate to *Hyakunin Isshu* using an existing speech synthesis function.

3) *Nazokake* Riddles

A *nazokake* riddle has the format "Given two words X and Y that are apparently different in meaning, what do they have in common? The answer is Z." In the "katarai" service, the user can choose one of four random words displayed by the system or input any word desired as theme X to enjoy a *nazokake* riddle. The construction of this game is based on text data collected from Social Networking Services (SNSs) using posts that have been converted to an "X, Y, Z" format consisting of three items of data.

---

*4 **my daiz member:** An extension function (service) that can be added to my daiz (see *5).

4) Quiz

We developed a quiz function as an interactive game that can be played as much as one likes. In this game, the system presents a random problem and four options as answers combining the correct answer and incorrect answers. The user may ask for a hint with respect to the given problem if needed. If the user answers correctly, the system assigns a score to that answer according to the difficulty of the problem.

In the "Word Play" member described below, total scores can be compared (ranked) among users by fixing the problems prepared in a certain period and tabulating the scores of each user.

3.2 Application Example

The "my daiz*5" speech dialogue agent offers "Word Play" as an interactive game using this technology (**Figure 3**). Although the main function of this agent is to provide information such as train service updates and weather, it includes an interactive game function with the aim of breaking down any psychological resistance that users may have about using agents and getting them to enjoy interaction with agents.

The "my daiz" member Word Play enables button-based operations by a Graphical User Interface (GUI)*6 as an original feature in addition to voice/text-based operations. This feature reduces input mistakes caused by speech-recognition errors,

Figure 3 Word Play

---

*5 **my daiz:** A speech dialogue agent that runs on smartphones and tablets, providing a wide range of information suited to the user.

*6 **GUI:** A superior type of interface that offers visibility and intuitive operability by expressing operations and display objects in the form of buttons, icons, etc.

typing errors, etc. In addition, the quiz function adds mechanisms for improving the user retention rate such as applying a limit to the number of answers that can be given per day and assigning a score according to the level of difficulty of each question. The *Hyakunin Isshu* function, meanwhile, can read out the first three lines of a poem using "my daiz" speech synthesis to enable the user to practice for competitive play and to provide a learning effect.

4. Katarai Slack Application

4.1 Background

Chat tools like Slack®*7 are attracting attention due to the recent nationwide trend toward teleworking. However, issues particular to remote working are coming to light such as "a decline in communication" since there are less opportunities for interacting and conversing in a face-to-face manner in real space. Under these conditions, we can expect dialogue that is also entertaining to stimulate person-to-person and person-to-thing communication. With this in mind, we developed a dialogue application that runs on Slack—a leading chat tool—to make it easy for users to enjoy entertaining dialogue. As an application that can be installed in Slack, it is the first to have a function for enlivening dialogue within a channel*8.

4.2 Functions

The following functions become available simply by installing this Slack application.

(1) Replies to posts

This function posts a natural reply to the content of a message posted on that channel. The text making up the reply uses output from the katarai chat-oriented dialogue engine. To prevent the channel log (record of messages) from flowing in a disordered manner, this function issues replies with a fixed probability instead of replying to every user utterance.

(2) Word play that picks up keywords in posts

If words stored in the database of *nazokake* riddles described above should happen to be included in messages posted on the channel, this function will reply with word play with a fixed probability. For example, given the post "I would love to have some ramen," the application might reply with "What do the words 'ramen' and 'Gion Festival' (one of the most famous festivals in Japan) have in common? The answer is *dashi* (which, in Japanese, can mean either 'soup stock' or 'parade float')."

(3) Automatic posting of *nazokake* riddles and quizzes

This function automatically posts content such as *nazokake* riddles to liven up a channel. This can prevent chat-oriented channels from becoming buried or inactive and induce users to make posts on such channels.

We administered a questionnaire consisting of eight items to nine users who participated in a trial of this application (**Figure 4**). Then, when using "4. Undecided" on a seven-point Likert scale*9 as baseline, we conducted a one sample t-test*10 against the evaluation mean for each question and found that a significant difference on a 5% level existed for items like "I wanted to send messages" and "It lowered barriers to sending messages."

---

*7 Slack®: A registered trademark of Slack Technologies, Inc.

*8 Channel: In Slack, a place for organizing and sharing messages in units of projects, teams, etc.

*9 Likert scale: A type of answer scale used in questionnaires and other types of psychological examinations. It enables the responder to indicate the extent to which he/she agrees with the statement presented. A Likert scale is generally a five-point scale but seven-point and nine-point scales are also used.

*10 One sample t-test: A statistical test that determines whether the mean of a population is equal to a specific value. In this article, we conducted this test under the hypothesis (null hypothesis) that the mean of questionnaire results is 4.0 (= specific value).

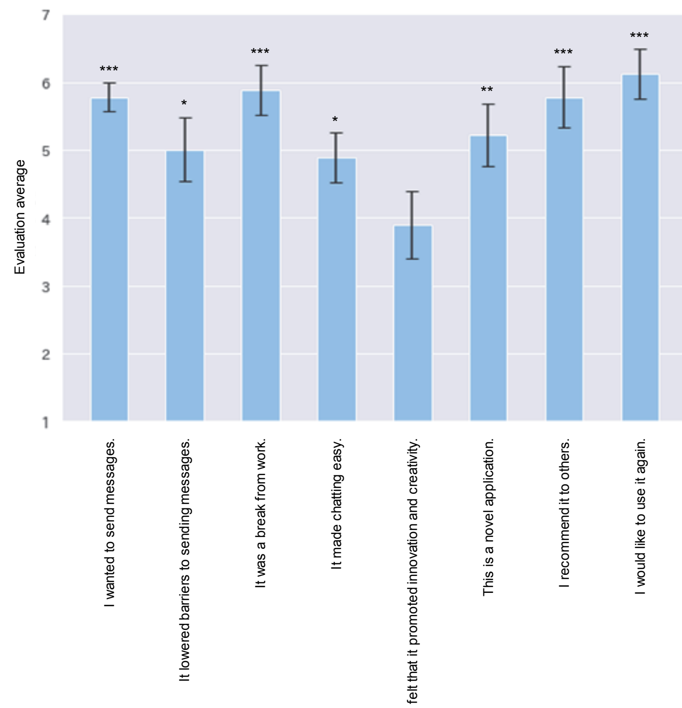


Figure 4 Result of questionnaire on katarai application

These results suggest that using this application should lower the psychological barriers to posting messages on a channel and eliminate a sense of isolation due to a drop in face-to-face interaction while also promoting communication within the organization. In the figure, the error bars indicate standard deviation (variation in collected data) and the number of asterisks (*) at the top of each bar graph indicate level of significance (*: $p < 0.05$; **: $p < 0.01$; *** $p < 0.005$).

5. Conclusion

In this article, we described character chat technology, word play technology, and an entertaining dialogue application. Character chat technology makes it possible to reproduce dialogue with any type of character by constructing a dialogue system based on utterance-pair data in the style of the desired character. Word play technology provides four types of games including "Yamanote Line" and "*nazokake* riddles" constructed on the basis of a knowledge base such as Wikipedia. These technologies make it possible to give chatbots an entertaining nature that makes them friendly and approachable to users. We also developed a Slack application that provides entertaining dialogue to make it easy for users to enjoy entertaining and friendly interaction.

Going forward, we plan to extract problems from actual services using these technologies and to make improvements toward services that can provide dialogue that is even friendlier and more entertaining to the user.

REFERENCES

[1] docomo Developer support website.
https://dev.smt.docomo.ne.jp/

[2] katarai website.
https://www.katar.ai/

[3] R. Higashinaka, M. Mizukami, H. Kawabata, E. Yamaguchi, N. Adachi and J. Tomita: "Role play-based question answering by real users for building chatbots with consistent personalities," Proc. of SIGDIAL, Jul. 2018.

Technology Reports (Special Articles) Efficient Bike Sharing Bike Demand Prediction Optimal Relocation

Special Articles on AI for Improving Services and Solving Social Problems

AI-based Optimization of Bike Relocation in Docomo Bike Share Service

X-Tech Development Department Shin Ishiguro Tomohiro Mimura
Satoshi Kawasaki
Service Innovation Department Yusuke Fukazawa†1
Smart-life Planning Department Atsuko Murozumi†2

The expanded use of bike-sharing services in recent years is not without its problems as reflected in particular by overflows or shortages of bikes at exclusive bike stations. In response to these issues, NTT DOCOMO has developed a bike-relocation optimization system for the pickup/drop-off trucks of the Docomo Bike Share service. This system predicts demand for bikes, optimizes the bike-relocation plan based on future predictions, and recommends relocation work to pickup/drop-off trucks. Using this system to relocate bikes is expected to alleviate overflows and shortages of bikes at each bike station and to maintain an easy-to-use bike-sharing environment for users.

1. Introduction

Bicycles are becoming increasingly popular in urban transportation or sightseeing as an easy means of traveling short distances from a train station to one's destination or between multiple destinations. In addition, the demand for bicycles has grown all the more since the beginning of the COVID-19 pandemic as they are a type of individual transportation that makes it easy to avoid the "Three Cs" (closed spaces, crowded places, and close-contact settings). For individuals, however, owning a bicycle for use in an urban environment is generally difficult due to the need for periodic contracts to reserve

†1 Currently General Affairs Department
†2 Currently Human Resources Management Department

space in a bicycle parking lot, the need to carry one's bike onto a train if traveling between cities, etc. Against this background, bike sharing has been attracting attention and its use has been increasing.

DOCOMO BIKE SHARE, INC. has been rolling out its bike-sharing business called Docomo Bike Share with the support of NTT DOCOMO. In this service, exclusive bike stations are set up at various locations in the city so that users can select bike stations of their choice to rent and return bikes. This service allows rented bikes to be returned to any bike station, i.e., a user need not return to the bike station where the bike was rented, which makes this service highly convenient for traveling short distances. The use of Docomo Bike Share continues to grow. Annual number of rentals increased

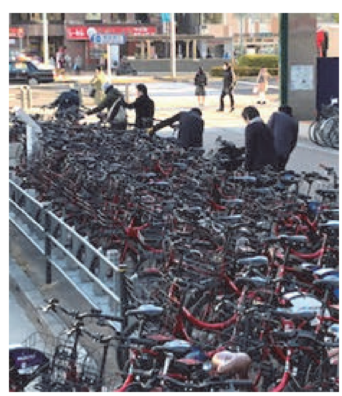
Photo 1　Motor-assist bike in Docomo Bike Share service

dramatically from about 40,000 in FY2011 to about 10,000,000 in FY2019. In light of this demand, DOCOMO BIKE SHARE, INC. has come to provide about 830 bike stations and 8,300 bikes in Tokyo as of 2019.

The Docomo Bike Share service features motor-assist bicycles (**Photo 1** [1]), which means an easy-to-use service even with the many hills that make up Japan's urban landscape. In addition, since bike maintenance and battery charging are also part of the service, users can enjoy the use of bikes that are always well maintained.

Bike sharing as typified by the Docomo Bike Share service is designed so that a user can travel freely from a certain bike station to another bike station, which means that a bike lent out to a user may be returned to a bike station different from the bike station where it was rented. As a result, the allocation of bikes may become locally unbalanced depending on bike usage. This, in turn, means that certain bike stations may be short of bikes preventing users from renting bikes or may be overflowing with bikes preventing users from returning their bikes (**Photo 2**). How to go about alleviating

(a) Bike shortage



(b) Bike overflow

Photo 2　Unbalanced conditions in bike allocation

this imbalance in bike allocation has become a common problem in the operation of bike-sharing services. In particular, the Docomo Bike Share service allows for bike parking other than on a rack at a bike station, so there are cases in which many bikes are returned and parked outside the bike-station area causing an overflow of bikes that can be highly inconvenient to property owners and pedestrians (Photo 2 (b)).

To alleviate this imbalance and allocate an appropriate number of bikes, the Docomo Bike Share service performs bike relocation using pickup/drop-off trucks. This relocation process eliminates an imbalance in bike numbers by loading bikes from bike stations with an overflow of bikes and carrying them to bike stations with a shortage of bikes. Bike-share use, however, continues to increase, so there is a need to deal with this imbalance in bike allocation by efficiently operating a limited number of pickup/drop-off trucks.

To meet this need, NTT DOCOMO developed a relocation optimization system that collects data on bike-share usage and local population statistics in real time and predicts the demand for renting and returning bikes at each bike station (**Figure 1**). In this way, the system can select which bike stations should be targeted for relocation work to alleviate future overflows or shortages and advise pickup/drop-off truck drivers the order in which those bike stations should be visited and the number of bikes that should be picked up and relocated. In this article, we describe our work in optimizing the bike relocation process in the Docomo Bike Share service.

2. Design of Docomo Bike Share Service and Rebalancing Problem

Each bike station provided by Docomo Bike Share incudes racks for docking bikes, but a rack itself is not equipped with a rental/return function but rather with a beacon for detecting bike location in the immediate vicinity. Specifically, short-range communication between the beacon and bike makes it possible to determine the state of bike return near the bike station. In this way, the service has been designed to enable bike status to be detected by a beacon so that the number of racks

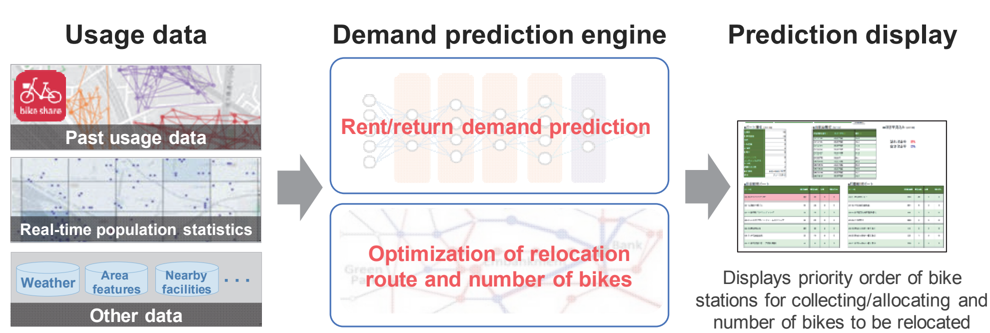


Figure 1　Overview of relocation optimization system

does not restrict the maximum number of bikes located at a bike station. In other words, the design is such that the number of bikes exceeding the number of racks cannot be physically limited. This design makes effective use of the space occupied by the bike station and enables many bikes to be parked, but if a large number of bikes happen to be returned, it can also result in a situation in which bikes overflow the parking area.

To evaluate bike overflow conditions at a bike station in the Docomo Bike Share service, the number of bikes that constitute an overflow limit is determined at each bike station on the basis of a threshold value that depends on the time period. This value serves as an indicator of an overflow situation.

In addition, if it is predicted that demand will exceed the number of available bikes at a bike station, opportunities for using the service may be lost. To evaluate such lost opportunities, the number of bikes that constitute a shortage limit is determined at each bike station for use as an indicator of a shortage situation. At most typical bike stations, the existence of demand for bike rentals in a state in which no bike is available is considered to be a lost opportunity, so this bike shortage limit is set to zero.

Here, a number of bikes at a bike station that is under the bike overflow limit and above the bike shortage limit is defined as an appropriate number of allocated bikes.

Renting or returning bikes in excess of these limits results in a state of bike overflow or shortage at a bike station as shown by the diagram in **Figure 2**.

Up to now, pickup/drop-off truck drivers in the Docomo Bike Share service have been performing bike relocation by judging future demand on their own based on past usage and weather forecasts. However, in the case of bike demand that can change from hour to hour, a manually prepared work plan cannot necessarily achieve optimal relocation.

With this in mind, we developed a relocation optimization system for making relocation work more efficient based on demand prediction. The aim of this system is to alleviate extreme bike overflow or shortage at bike stations by predicting future

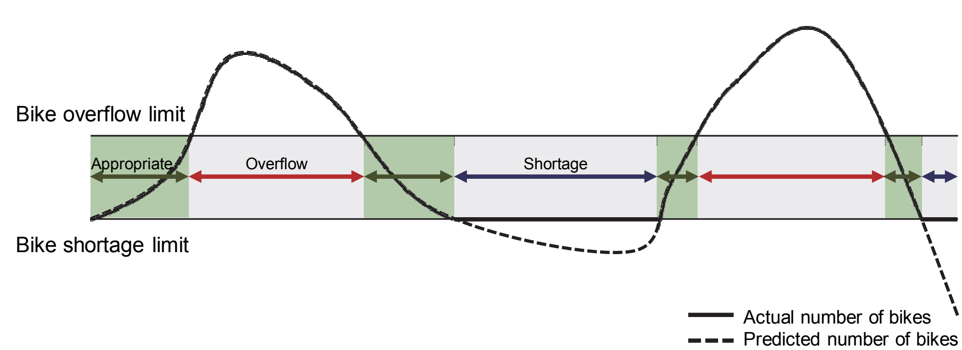


Figure 2　Diagram of bike overflow and shortage at bike station

overflow and shortage at each bike station, planning a procedure for allocating and collecting bikes so as to balance out the number of bikes at each bike station, and conveying that information to pickup/drop-off truck drivers as a basis for performing relocation work.

3. Overview of Relocation Optimization System

3.1 Demand Prediction Function

The renting and returning of bikes in a bike sharing service are affected by a variety of factors such as the number of people moving about, time of day, day of the week, weather, and the holding of events. We therefore created a model of demand prediction to reflect such ever-changing demand. For this model, we used, in particular, eXtreme Gradient Boosting (XGBoost)*1 and an extension of a time-series deep learning technique*2 [2].

An overview of the demand prediction function is shown in **Figure 3**. This function predicts as objective variables the number of rented bikes and number of returned bikes that occur over a period of one hour for each of the next 24 hours. It creates hourly models equivalent to a period of 24 hours for both rentals and returns resulting in a total of 48 models and calculates the number of rented bikes and number of returned bikes for each time period as prediction results. Consequently, by determining the difference between the number of rented bikes and the number of returned bikes in each hour, it becomes possible to calculate the amount of change in the number of available bikes at that time. This change in the number of available bikes can then be used to calculate the predicted values

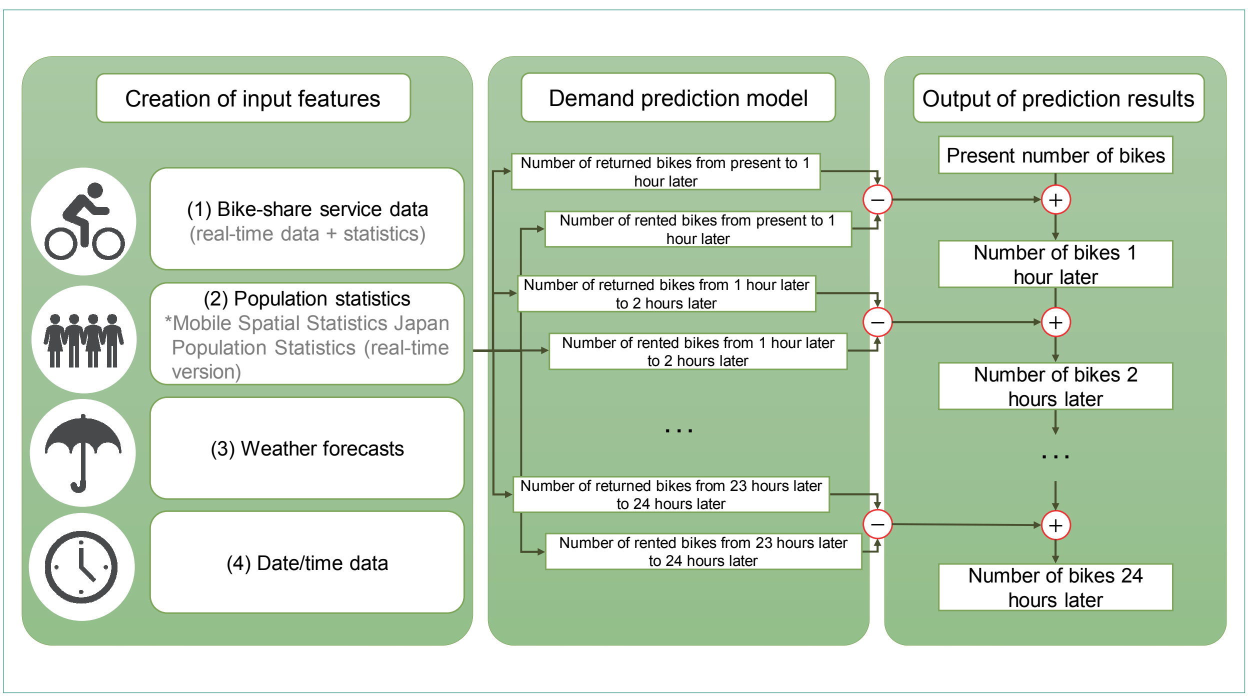


Figure 3　Overview of demand prediction function

*1　XGBoost: A machine learning technique that achieves high-accuracy regressive prediction by taking a majority vote based on a prediction model using different decision trees.

*2　Deep learning technique: A machine learning technique that achieves regressive prediction taking time-series changes into account through a recurrent neural network.

for the number of available bikes in each hour by determining the cumulative sum of available bikes taking the number at the present time as the initial value.

The input features used for learning are (1) bike-share service data, (2) population statistics, (3) weather forecasts, and (4) data/time data. Using these features in model learning should make it possible to mechanically determine any correlation between change in future demand and population, weather, and past bike-share usage and to achieve more accurate predictions of future demand.

(1) Bike-share service data includes the number of rented bikes and number of returned bikes up to that hour at each bike station and the amount of change in those values. It also includes the average values for the number of rented bikes and number of returned bikes for the same day of the week and time period in the past as well as the average value of that change as statistical quantities.

(2) Population statistics refers to the population within 500 m of each bike station and amount of change. We expect the use of real-time data here to enable change in human movement to be grasped particularly in predictions targeting the next one to two hours.

(3) Weather forecasts consist of the latest weather data on rainfall, wind, temperature, humidity, and atmospheric pressure for the hour targeted for prediction.

(4) Date/time data consists of the time, year, month, date, and day of the week for the hour targeted for prediction.

3.2 Relocation Planning Function

The relocation planning function shown in **Figure 4** is described in detail below.

1) Inference of Bike Overflow/shortage Based on Prediction Results

A future prediction value of the number of available bikes can be obtained by the demand prediction function described above. The number of available bikes for each hour as determined by this function is the cumulative sum of the change in the number of available bikes, and as shown in Fig. 2, the prediction value may take on a negative value. Now, if the prediction value of the number of bikes for each hour at each bike station should rise above the bike overflow limit established for each bike station, it could be inferred that the bike station is overflowing by the number of bikes exceeding that limit. Conversely, if the prediction value should drop below the bike shortage limit, it could be inferred that the bike station is short by the number of bikes exceeding that limit

2) Recommended Work Based on Evaluation Values

The next step is to formulate a work plan for collecting and allocating bikes for these amounts of overflow and shortage and to then dispatch pickup/drop-off trucks to do this work. At this time, the picking up of excess bikes for relocation and the dropping off of bikes at a bike station having a shortage of bikes constitute a sequence of tasks, so bike collecting and bike allocating are an inseparable pair of tasks. Additionally, since the truck moves during the relocation process, the distance from the truck's present location to the bike station targeted for collecting bikes and the distance from that bike station to the bike station targeted for allocating bikes can be calculated. Taking

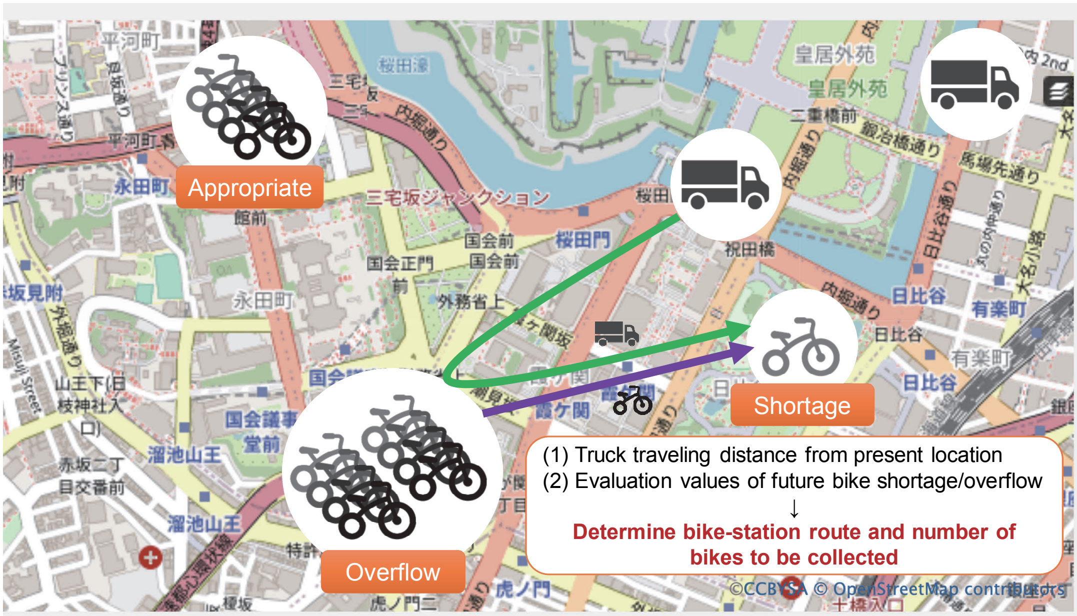


Figure 4　Overview of relocation planning function

these distances into account, an overflow evaluation value and a shortage evaluation value can be calculated for each bike station. The overflow evaluation value takes on a larger value as the number of excess bikes becomes larger and the distance becomes shorter. In such a case, the bike station would be evaluated as having high collection priority. The shortage evaluation value, in turn, takes on a larger value as the number of deficient bikes becomes larger and the distance becomes shorter. In this case, the bike station would be evaluated as having high allocation priority. Each of these evaluation values can change from hour to hour due to the time taken up by relocation work and truck movement and to fluctuation in the supply and demand for bicycles. These evaluation values can therefore be recalculated based on a request received from the truck on completing its work. At this time, the bike station receiving the highest evaluation value is given priority and the next order of tasks is determined.

3) Determining the Number of Bikes to be Collected/allocated Based on Evaluation Values

The number of bikes to be collected or allocated at a bike station selected on the basis of the corresponding evaluation value is as follows. In collection work, as many bikes as needed to bring the number of available bikes below the bike overflow limit will be picked up, and in relocation work, as many bikes as needed to bring the number of available bikes above the bike shortage limit will be allocated.

Moreover, due to the work of collecting and allocating bikes, the number of bikes currently loaded on the truck changes. For this reason, the following values are used as upper limits to the number of bikes to be collected and the number of bikes to be allocated.

Bikes to be collected = min (number of excess bikes, max. number of loadable bikes – current number of loaded bikes)

Bikes to be allocated = min (number of deficient bikes, current number of loaded bikes)

4) Determining a Bike Relocation Plan

This step determines work priority for each time period based on the evaluation values calculated for each bike station and the number of bikes required for collection and allocation. It extracts the bike-station pair having the highest evaluation values for collection and allocation work and creates a plan for collecting and allocating bikes between those bike stations with the maximum number of bikes that can presently be loaded on the truck. Moreover, in the case that the number of bikes to be collected and the number of bikes to be allocated differ (that is, either the number of bikes to be collected or the number of bikes to be allocated is smaller), collection and allocation work will continue for two more bike stations to maintain a balance between the number of bikes collected and allocated. When relocating bikes using multiple pickup/drop-off trucks, the possibility arises that the same bike station will be selected as a target for bike relocation in the same time period. To avoid this situation, a relocation plan for multiple trucks will be created when one truck is already working by selecting other bike-station pairs for other trucks.

4. Trial

We conducted a trial to assess the effectiveness of the demand-prediction and relocation-planning system described above. For this trial, we created a web application that presents the number of bikes to collect and relocate from one bike station to another in each time period via a tablet-type user interface and promoted its use by pickup/drop-off truck drivers. The target districts selected for the trial were Koto city, Minato city, Chuo city, Shinjuku city, and Shinagawa city in the Tokyo area. The trial was conducted in each of these wards for the periods listed in **Table 1**. Each of the pickup/drop-off trucks used in the experiment could load up to 30 bikes. Considering that the relocation of bikes throughout the Tokyo area could be quite complex, we decided to use a different number of trucks in each ward and performed relocation planning within each ward for that number of trucks. The number of trucks used in each ward is listed in **Table 2**. For learning data, we used bike-share service data, Mobile Spatial Statistics Japan Population Statistics (real-time version)*3, and Meso-Scale Model (MSM)*4 Grid Point Value (GPV) data of the Japan

Table 1 Trial period in each ward

Ward	Period
Koto, Minato	2019/11/1～15
Chuo, Shinjuku, Shinagawa	2019/11/1～20 2019/12/3～6

Table 2 Number of trucks used in each ward

Ward	Number of Trucks
Koto	3
Minato	2
Chuo	1
Shinjuku	1
Shinagawa	1

*3 Mobile Spatial Statistics® Japan Population Distribution Statistics (real-time version): Data obtained by calculating population distribution throughout Japan at 500 m and 10 minute intervals for up to 1 hour to 30 minutes before through statistical analysis using mobile data. Mobile Spatial Statistics is a registered trademark of NTT DOCOMO, Inc.

*4 MSM: A weather forecasting model introduced by the Japan Meteorological Agency (JMA) that makes predictions once every 3 hours, 39 hours into the future.

Meteorological Business Support Center and set the learning period from October 2017 to March 2019. Details on these data are given in **Table 3**.

For relocation by the proposed system, **Figure 5** shows the results of comparing evaluation values for the number of deficient bikes (shortage) and number of excess bikes (overflow) in periods before and after the trial. Before the trial, relocation was performed based on the judgment of the pickup/drop-off truck drivers, and in the period after the trial, it was performed based on the instructions received from this system. It was found that applying this system could improve shortage and overflow conditions in all of the five Tokyo wards targeted by this trial. It could also be seen that percentage improvements in shortage and overflow conditions differed from one ward to another.

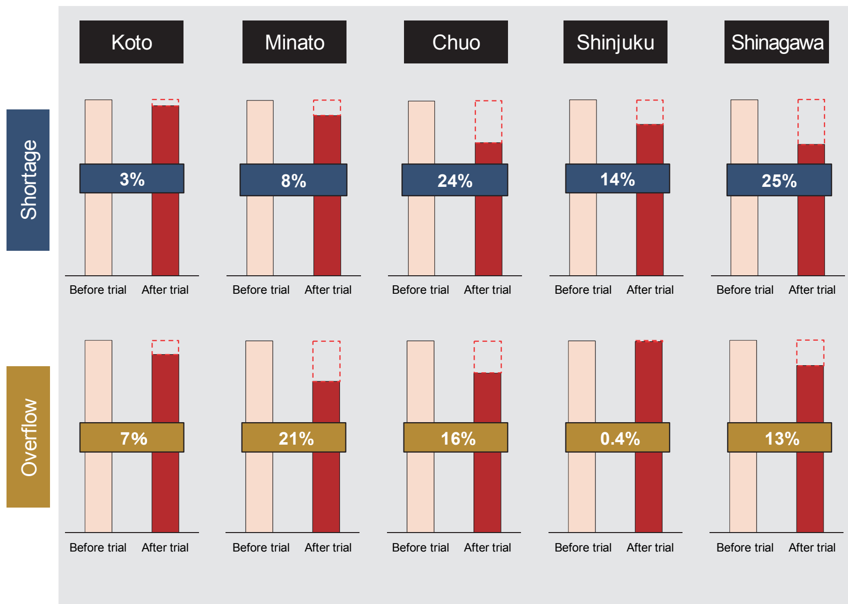
In Shinjuku city, the improvement effect with respect to the number of excess bikes was only 0.4%. In this regard, work based on pickup/drop-off trucks is often efficient if more bikes than usual can be moved in one relocation operation. Shinjuku city, however, features many small-scale bike stations each of which allocate only a small number of bikes, so achieving an improvement in overflow requires the patrolling of many bike stations in the relocation process. We consider that this is why the efficiency-improvement effect by this system turned out to be small for relocation by pickup/drop-off trucks in this ward.

Other wards, however, include large-scale bike stations that can allocate many bikes. This made it possible to perform bike relocation making good use of the features of pickup/drop-off trucks and to achieve improvement effects in overflow and shortage conditions.

We next describe an experiment related to prediction accuracy at the time of an event. In bike sharing, most usage is periodic in nature corresponding to the days of the week in which people

Table 3　Details of input features

Name of Data	Details
Bike-share service data	Approx. 5,900 bikes
	Bike stations: Approx. 580
	Approx. 7,300,000 bike rentals
	Each of these figures represent total amounts for all Tokyo areas as of March 2019
Population statistics *Mobile Spatial Statistics Japan Population Statistics (real-time version)	Mobile phones Approx. 78,450,000 (excluding corporate and MVNO phones)
	Total number of phones for all of Japan as of March 2019. Uses statistics tabulated in mesh form
	500 m mesh, once every 10 min
Weather forecasts *Meso-Scale Model (MSM) GPV	Rainfall, temperature, humidity, wind, wind direction, atmospheric pressure, and cloudiness
	Issues forecasts once every 3 hours, 39 hours into the future
	1,000 m mesh



*Values overlapping each pair of bar graphs indicates percentage improvement in shortage/overflow evaluation values

Figure 5 Comparing improvement effect in shortage and overflow of bikes before and after trial

commute by train, time of day, etc. This type of demand can be predicted with sufficient accuracy by statistically analyzing bike-share service data. On the other hand, the occurrence of special demand due to a major event cannot be explained simply by analyzing the periodicity of bike-share usage and is consequently difficult to predict. For this reason, adding population statistics to input data enables the system to obtain changes in the neighboring population and to predict change in usage demand at the time of an event. The results of predicting change in bike usage at a bike station in the neighborhood of a certain event venue are shown in **Figure 6**. These results show that using population statistics could improve the accuracy of predictions in relation to times of peak demand for some bike stations. They also show that real-time population statistics can serve to explain changes in demand due to an increase/decrease in the neighboring population at places such as event venues where amount and time of people movement can easily change depending on the day. We therefore consider that real-time population statistics can help improve prediction accuracy.

5. Conclusion

This article described our efforts in optimizing the relocation of bikes for improving operation of the Docomo Bike Share service. We developed, in particular, a system that combines a demand prediction function and a relocation planning function to predict imbalance in future demand and supply and to propose bike relocation according to current

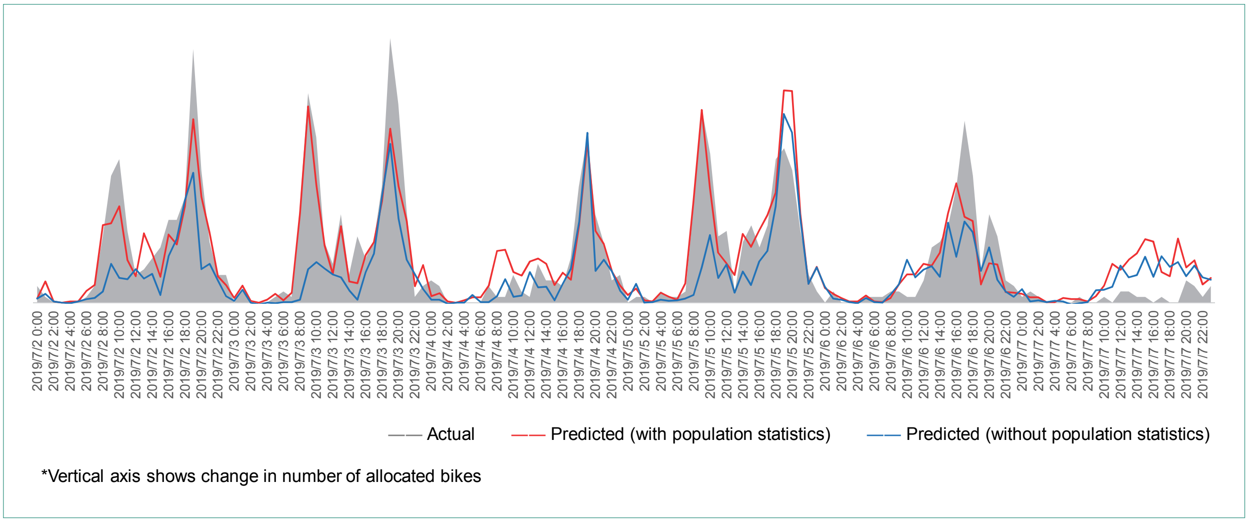


Figure 6 Example of bike station with improved prediction (July 2 – July 7, 2019)

conditions. We showed that relocating bikes with this system reduced the number of bike stations in an overflow or shortage state and made the work of relocation more efficient in five districts within the Tokyo area. We also found that the use of population statistics makes it possible to estimate fluctuations in demand and improve the accuracy of predictions at locations such as event venues where making predictions solely on the basis of past bike share usage is difficult. Going forward, we plan to study the use of this bike relocation optimization system throughout Japan and the optimization of other means of transportation by applying the know-how obtained in the development of this system.

REFERENCES

[1] NTT DOCOMO Bike Share Service.
https://docomo-cycle.jp/?lang=en

[2] T. Mimura, S. Ishiguro, S. Kawasaki and Y. Fukazawa: "Bike-share demand prediction using attention based sequence to sequence and conditional variational autoencoder," In Proc. of the 3rd ACM SIGSPATIAL International Workshop on Prediction of Human Mobility, pp.41–44, Nov. 2019.

Technology Reports (Special Articles) Mental Health Care Stress Cognitive Performance

Special Articles on AI for Improving Services and Solving Social Problems

Smartphone Log-based Stress and Cognitive Performance Estimation Technology

X-Tech Development Department Naoki Yamamoto Takashi Hamatani
Keiichi Ochiai
Service Innovation Department Shigeki Tanaka Yusuke Fukazawa†

Around the world, the number of patients with mental disorders is increasing every year, therefore prevention of deterioration of mental health has become a global problem. However, it is difficult to recognize one's own mental health state correctly because we cannot confirm it visually and objectively in daily life. To tackle this problem, we propose a method that estimates the mental state in terms of stress level and cognitive performance by extracting behavioral features from a smartphone log via passive sensing.

This technology makes it possible to visualize the user's mental health state and support self-care by encouraging users to be aware of their own stress conditions.

1. Introduction

The number of patients with mental disorders is increasing every year, and there is growing social interest in mental health care. According to the World Health Organization (WHO), more than 300 million people worldwide suffer from depression, and more than 90% of suicides are reported to be caused by mental illness [1]. Excessive stress is a known factor of mental disorders. Thus, to maintain a healthy mental state, it is important to help people to be aware of the stress they face and

† Currently, General Affairs Department

enable self-prevention, although it is difficult for people to recognize their own stress state correctly.

To address this, there have been a number of efforts to estimate the mental health of users based on logs obtained from smartphones and wearable devices [2]. Since smartphones are owned, carried and used by individuals at all times, smartphone logs reflect the daily activities of users. Assuming that the user's activity and mental health states influence each other, the mental health state of the user should be reflected in smartphone logs. Since the logs of daily smartphone usage are collected, the user does not need to perform any special operations, and the possibility that the user will be burdened is minimized. Additionally, smartphone logs are collected constantly, making it possible to continuously assess the user's mental health state.

Although stress is considered as a risk factor in mental health deterioration, moderate stress has the potential to tighten the mind and improve the efficiency of work and study [3]. However, a state of high stress and low efficiency in work and study means that the user's mental health is deteriorating. Therefore, even in high-stress states, there are situations that affect users positively and situations that affect them negatively, so it is necessary to separate and interpret the good and bad aspects of high-stress states.

In this study, in addition to estimating the degree of daily stress based on smartphone logs, we aimed to identify whether the stress a person is facing is appropriate from the perspective of the cognitive performance (concentration) by estimating the cognitive performance (response speed and accuracy of judgment), which indicates the ability for intellectual activity.

In this article, we describe the details of our proposed method and the results of demonstration experiments using the proposed method.

2. Proposed Model

Figure 1 shows an overview of the proposed model. In the proposed method, we constructed an estimation model with the smartphone log as an explanatory variable, a stress index based on heart rate data, and an cognitive performance score based on the cognitive performance measurement task (Go/NoGo task) as the respective objective variables. The stress index and cognitive performance score are used only when building the estimation model. After the model is built, estimation of stress and cognitive performance states can be done simply by collecting smartphone logs.

In this study, we collected a dataset from volunteer employees working in the R&D division of NTT DOCOMO INC. Thirty-nine employees ranging from their 20s to 50s, consisting of 34 males and 5 females, participated in the data collection. The duration of the study was three months from November 2017 to January 2018, and study participants were asked to cooperate for up to 42 days during this period.

The study was approved by the ethics committee of the Graduate School of Medicine, part of the Faculty of Medicine at the University of Tokyo (examination number: 2017-001).

2.1 Behavioral Feature Values

The smartphone logs collected and the behavioral feature values used in this study are shown in **Table 1**. A total of 1,349 days of smartphone logs

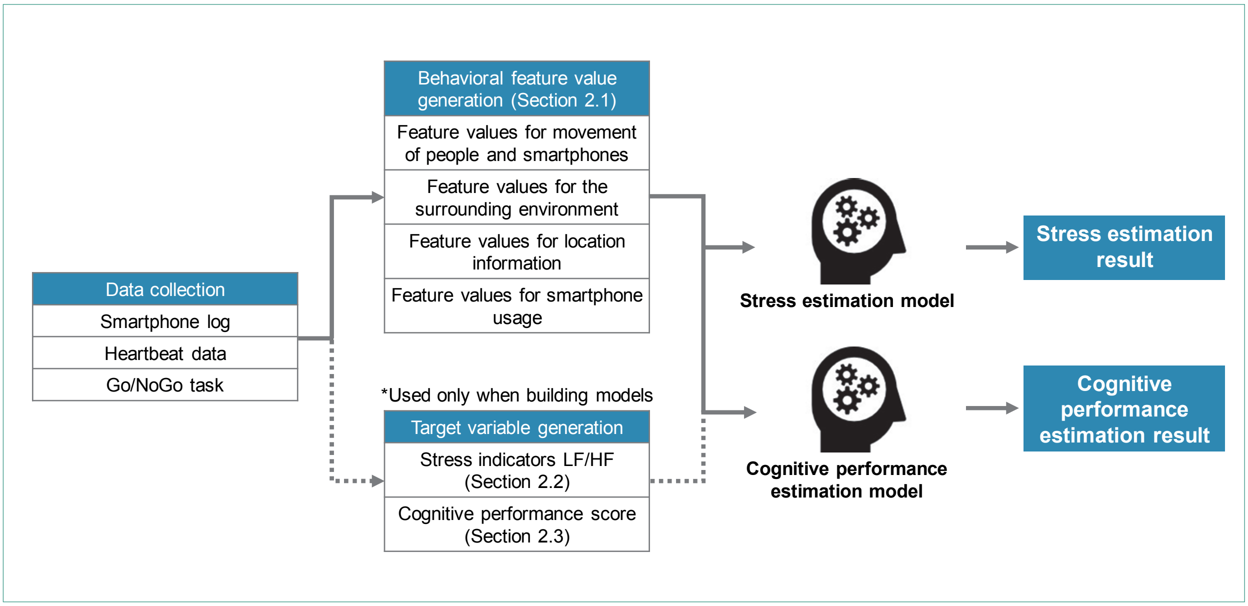


Figure 1　Stress, cognitive performance estimation model

Table 1　Behavioral feature values of stress, cognitive performance estimation model

Smartphone log	Generated feature values (excerpt)
3-axis Accelerometer Gyro sensor	Mean, variance, max, min, max-min; magnitude of vector; correlation coefficients for each axis
Ambient illumination Air pressure Battery level	Mean, variance, max, min, max-min
GPS	Max/min/max-min of latitude/longitude/altitude; the total distance traveled; the number of places visited
Earbud connections Charger attached	The number of connections; connected time per day
Screen ON/OFF	The number of times; mean, variance, max, min of screen-on times
App usage history	The number of used applications by category (games, communication tools, etc.)
Activity Recognition API	Percentage of each state (vehicle, bicycle, on foot, running, walking, still, tilting, unknown, ride, move)
Storage usage amount	Value of remaining capacity
Day of the week	Monday～Sunday; holiday

were acquired through the data collection test.

Behavioral feature values were generated under the following four perspectives.

(1) Feature values for movement of people and smartphones

We assumed that the user's mental state

would be reflected in their behavior. For example, when people feel stress they are likely to handle their smartphones more roughly.

(2) Feature values for the surrounding environment

We assumed that changes in the surrounding environment would affect the user's mental state and behavior. For example, the user may feel stress due to sudden changes in air pressure.

(3) Feature values for location information

We assumed that the user's mental state would affect how they traveled and where they visited. We also assumed that places the user visited would affect the user's mental state. For example, the user may feel stress if the total distance traveled or the number of places visited in a day becomes greater than usual.

(4) Feature values for smartphone operation

We assumed that the usage pattern and purpose of smartphones would change depending on the user's mental state. For example, if the user is feeling stressed about interpersonal relationships, the number of times he or she uses communication applications may decrease.

We estimated the daily stress and cognitive performance of users based on these feature values.

2.2 Stress Index

In this study, we collected 639 days of ElectroCardioGram (ECG) data sets using heartbeat sensor and calculated the physiological index LF/HF [4] based on the ratio of Low Frequency (LF) to High Frequency (HF)*1 components as the stress index. LF and HF represent the activity of sympathetic and parasympathetic nerves, respectively. The ratio of LF and HF is a known index that represents stress. A higher value of LF/HF means stress is felt.

In this study, we estimate the stress on a given day as either higher or lower than its standard value.

First, the daily average values of LF/HF were calculated for each day, and then the average values for all those days were calculated as the standard values for each user. Next, we labeled each day's stress state by setting high stress days as days when the daily average value of LF/HF was above the standard value and low stress days as days when the daily average value of LF/HF was below the standard value.

2.3 Cognitive Performance Index

We used the Go/NoGo task as a method to measure users' cognitive performance [5]. The Go/NoGo task application implemented on Android devices is shown in **Figure 2**.

In this task, a letter of the alphabet is displayed on the smartphone screen continuously, and the user is required to tap the screen quickly (the Go reaction) only when one of the specified letters is displayed, and not to tap the screen (the NoGo reaction) when any other letter is displayed. The displayed letter is randomly selected from eight predetermined letters to be displayed, and letters are displayed a total of 72 times in approximately one minute. The sum of the number of correct Go and NoGo responses divided by 72 is the cognitive performance score in this task.

---

*1 Low-Frequency (LF) and High-Frequency (HF) components: Heartbeats are known to fluctuate periodically rather than always beat at a regular interval. These are the two main components of this fluctuation.

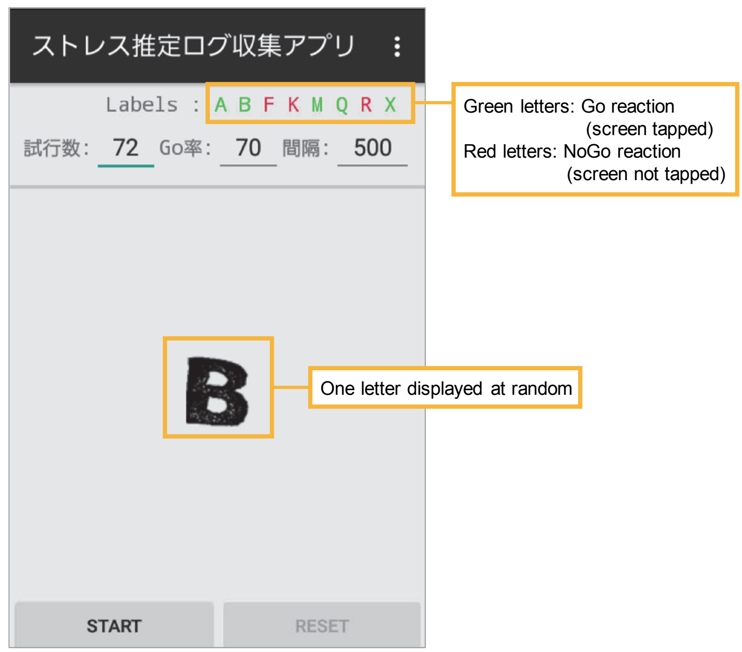


Figure 2　Go/NoGo task application

This task was performed three times a day, and the average of the scores was used as the representative value for the day. Data sets of 779 days of cognitive performance measurements were acquired through data collection test. As with stress, to label each day with a cognitive performance score, the average cognitive performance score for each user was calculated as the standard value, and days with an cognitive performance representative value above the standard value were set as high cognitive performance. and days with a value below the standard value were set as low cognitive performance.

2.4 Performance Evaluation

The estimation accuracy of stress and cognitive performance models was evaluated based on 1,349 days of smartphone logs, data sets of 639 days of ECG data and 779 days of cognitive performance measurement results. The numbers of these three types of data did not match because the wearing of the chest electrocardiograph and the implementation of the Go/NoGo task application were done voluntarily by the study participants to the extent that they did not affect their daily lives. Therefore, the performance evaluation was conducted using a total of 554 days of data, through analysis of days on which all three types of data (smartphone logs, ECG data sets, and cognitive performance measurement results data sets) were available.

We built two models to estimate stress and cognitive performance using a machine learning algorithm with the behavioral feature values as explanatory variables and stress and cognitive performance state labels as objective variables.

In evaluation, we defined one user as the target user for model evaluation and built a model based on the data sets of other users to evaluate the performance of the model for unknown users. We evaluated estimation result correctness by comparing the estimation result of the target user with the stress index based on ECG data sets and the Go/NoGo task of the target user. We evaluated each user assuming that the user is unknown and combined the evaluation results of all users to form the final evaluation results. We used three indices to evaluate the estimated results: Accuracy*2, Sensitivity*3 and Specificity*4.

Table 2 shows the estimation accuracy. The proposed method achieved an accuracy of more than 0.700 for both stress and cognitive performance. However, the sensitivity of both stress and cognitive performance was low compared to the specificity, and there is room for technical improvement to accurately estimate the high stress and high cognitive performance states.

3. External Verification Testing

3.1 Development of a Verification Application

We developed an experimental application that incorporates the stress estimation model and the cognitive performance estimation model. **Figure 3** shows the estimation results presentation screen of the application. This application constantly collects smartphone logs and stores the logs in internal storage. The stored logs are sent to the server once a day, and stress and cognitive performance are estimated on the server. The estimation results are sent to the smartphone and presented to the user through the application.

3.2 Demonstration Experiment in the METI Living Lab

1) Overview of the METI Living Lab*5

From December 2019 to March 2020, NTT DOCOMO INC. participated in the Living Lab sponsored by the Ministry of Economy, Trade and Industry (METI) (2019 Survey on Small and Medium Enterprises (Survey on the Creation of Innovative Social Problem-Solving Services in the Living Laboratory), hereinafter referred to as “Living Lab”). Based on the METI assumption that there is a causal relationship between stress and sleep, the Living Lab conducted an intervention program to improve the quality of sleep for METI employees

Table 2 Stress, cognitive performance estimation accuracy

	Accuracy	Sensitivity	Specificity
Stress	0.711	0.608	0.773
Cognitive performance	0.791	0.750	0.818

*2 Accuracy: In this article, this indicates the percentage of the total sample that is correctly classified as high stress (cognitive performance) or low stress (cognitive performance).

*3 Sensitivity: In this article, this indicates the percentage of users who should be assumed to be high stress (cognitive performance) who are correctly classified as high stress (cognitive performance).

*4 Specificity: In this article, this indicates the percentage of users who should be assumed to be low stress (cognitive performance) who are correctly classified as low stress (cognitive performance).

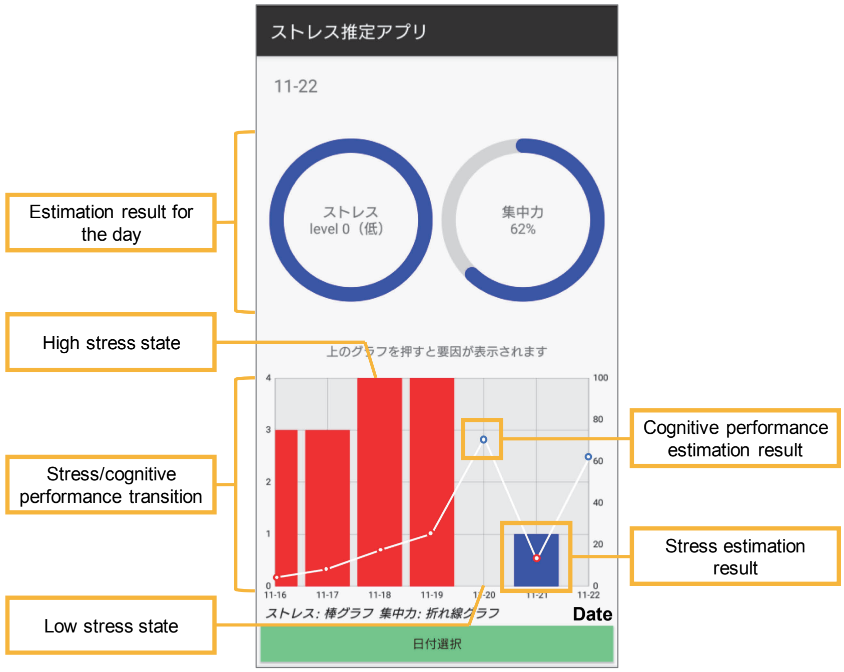


Figure 3　Stress, cognitive performance estimation application

(healthy adults) and verified the effects of the program on stress. Kodomo Mirai, Inc. was responsible for intervention and provided face-to-face and online guidance to improve sleep during the Living Lab period. In addition, a questionnaire was used to measure the stress state of participants at the beginning and end of the Living Lab.

NTT DOCOMO INC. provided the above experimental application as an instrument to measure the effectiveness of the intervention. The verification application was installed on the devices of 26 Living Lab participants. Then, 14 participants were asked to implement the sleep improvement program in accordance with the intervention, while the remaining 12 participants went about their daily lives as usual. During Living Lab, the application was used to estimate participants' daily stress and cognitive performance states.

2) Verification Results

Stress estimation results were evaluated from the following two perspectives.

(1) Intervention effectiveness

The stress states of the intervention group and the non-intervention group were compared to confirm whether there were any improvements in stress states.

(2) Validity of stress estimation results

We confirmed the degree of agreement between the results of the questionnaire at the beginning and end of Living Lab and the stress estimation results, and verified that it is possible to actually use the estimation

*5 Living Lab: An initiative to create services and products based on the knowledge gained by repeatedly conducting experiments and evaluations in a real environment in which users and providers work together to solve social issues.

model.

In the case of (1), we conducted the following analysis assuming that stress states improved with the implementation of the intervention. Periods i and ii refer to the periods before and after the intervention, respectively.

(A) Comparison of stress states between the intervention group (period i) and the non-intervention group (period i)

(B) Comparison of stress states between the intervention group (period ii) and the non-intervention group (period ii)

In this case, the assumption may be valid when there is no significant difference in (A) and a significant difference can be confirmed in (B). We used the Mann-Whitney U test*6 as the testing method, and set the significance level to 0.05.

Figure 4 shows the analysis results. The "*" in the figure indicates that the significance level has been met. In (A), we could not confirm any significant difference in the results of stress and cognitive performance estimation between the two groups. However, in (B), a significant difference between the two groups in the results of stress and cognitive performance estimation was confirmed, suggesting the intervention may have contributed to the improvement of stress states. In addition, the intervention group showed an increase in low stress

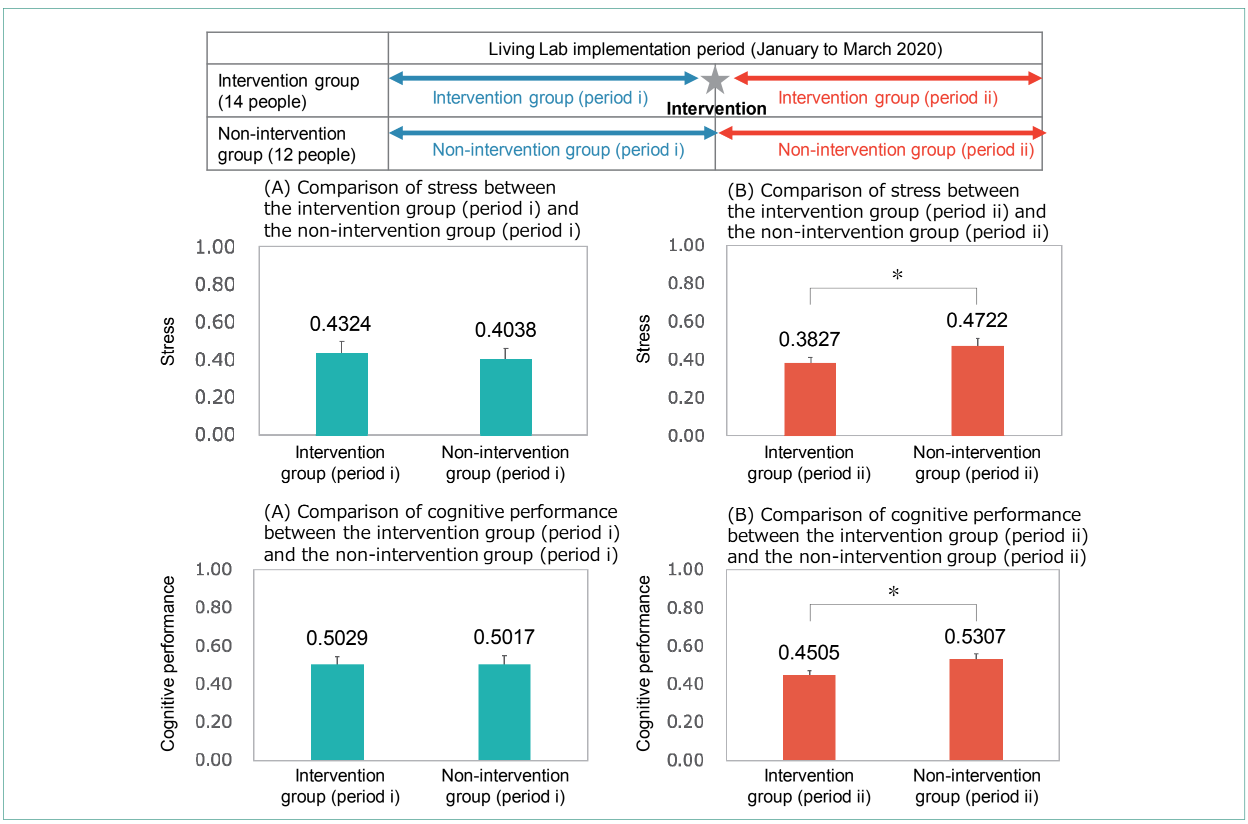


Figure 4 Changes in stress, cognitive performance due to the intervention

*6 Mann-Whitney U test: A non-parametric test for examining whether there is a statistically meaningful (significant) difference between two groups of data.

and low cognitive performance states, which can be interpreted as an increase in the relaxed state due to the intervention, which was intended to improve sleep quality.

Regarding (2), the questionnaire showed that the intervention group improved for the three items of stress response, subjective sleep quality and work results. From these results, it can be interpreted that the estimates made by the stress estimation model were reasonable. In addition, as mentioned above, we confirmed an increase in the relaxed state from the results of stress and cognitive performance estimation, while the questionnaire showed an improvement in work results. This can be interpreted as results of the intervention, which created a balance between the states of maintaining high work efficiency while feeling moderate stress, and relaxation.

4. Conclusion

In response to the growing interest in mental health care, we proposed a model for estimating the daily mental health of users based on logs collected from smartphones and described the results of a validation experiment using the model. Focusing on the advancement of the proposed technology by further verifying the results of this study, NTT DOCOMO INC. will accelerate efforts toward the practical application of the technology to support the mental health care of its customers. Going forward, we aim to provide comprehensive support for customers' physical and mental health care and contribute to extending their healthy life expectancy*7.

REFERENCES

[1] World Health Organization: "Preventing suicide: a global imperative," 2017.
http://www.who.int/mental health/suicide-prevention/world report 2014/en/

[2] Y. Fukazawa, N. Yamamoto, T. Hamatani, K. Ochiai, A. Uchiyama and K. Ohta: "Smartphone-based Mental State Estimation: A Survey from a Machine Learning Perspective," Journal of Information Processing, Vol.28, pp.16-30, 2020.

[3] C. Sandi: "Stress and cognition," Wiley Interdisciplinary Reviews Cognitive Science, Vol.4, No.3, pp.245-261, 2013.

[4] M. Vollmer: "A robust, simple and reliable measure of heart rate variability using relative RR intervals," 2015 Computing in Cardiology Conference (CinC), IEEE, pp.609-612, 2015.

[5] S. Hirose, J. Chikazoe, T. Watanabe, K. Jimura, A. Kunimatsu, O. Abe, K. Ohtomo, Y. Miyashita and S. Konishi: "Efficiency of go/no-go task performance implemented in the left hemisphere," Journal of Neuroscience, Vol.32, No.26, pp.9059-9065, 2012.

*7 Healthy life expectancy: The period of life during which a person is able to spend his or her daily life without any physical or mental health problems.

Technology Reports XR AR/MR Cloud Cyber-physical Fusion

AR/MR Cloud Technology to Provide Shared AR/MR Experiences across Multiple Devices

Communication Device Development Department Kouki Hayashi Osamu Goto Shinji Kimura Taishi Yamamoto

One of the new services expected to emerge in the 5G era is the "AR/MR Cloud". This article describes NTT DOCOMO's R&D efforts in the three major functions - the self-localization function, the spatial 3D model generation/management function, and the content space management function - necessary to realize the AR/MR Cloud that provides shared AR/MR experiences across multiple devices.

1. Introduction

NTT DOCOMO is working toward the realization of "cyber-physical fusion*1", which aims to improve and optimize the value of services and lifestyles in real space by converting information on people, things, and events in real space into data and predicting the future by utilizing AI technology on the collected data. Cyber-physical fusion refers to a world in which information about people, things, and events in real space is acquired through various sensors, and the information is aggregated in digital space in the cloud using communications such as 5th generation mobile communication systems (5G) to create digital twins*2. XR*3 are promising technologies for providing new services in the 5G era and more realistic experiences of the world of cyber-physical fusion. XR is a generic term for technologies such as Virtual Reality (VR)*4, Augmented Reality (AR)*5, and Mixed Reality (MR)*6 that provide new experiences by fusing virtual space with real space. XR technologies enable new experiences peering into the world of digital twins using Head-Mounted Displays (HMD)

*1 Cyber-physical fusion: Services and systems for realizing a better and more advanced society by collecting information in real space (physical space) from various sensors, etc. and linking it to virtual space (cyber space).

*2 Digital twin: A real-time reproduction in digital virtual space of the position, shape and various sensor information of an object in the real world.

such as glasses-type AR/MR devices, VR goggles or smartphones. Notable among these and attracting great attention in recent years are AR/MR Cloud technologies, which enable interactive and shared AR/MR experiences across multiple devices by superimposing cyberspace created with digital twins by aggregating various sensor information into real space.

"AR/MR Cloud" refers to technological infrastructure for superimposing AR/MR content on real space and enabling interactive sharing of AR/MR content across multiple devices. Services achievable with AR/MR Cloud technologies include:

- Highly immersive and interactive games such as virtually painting buildings in real space
- E-commerce that enables users to virtually draw furniture in a real room to check the fit before buying actual furniture or other items
- Advertising services that display advertisements and coupons, etc. that match the user's tastes and preferences or situation on storefronts, building walls, etc. as the user walks around a town
- Bulletin board services that enable users to freely write comments on the walls of buildings

At "DOCOMO Open House 2020 - Dawn of the 5G era and the Future Beyond -" held at Tokyo International Exhibition Center Aomi Exhibition Hall in January 2020, we held a demonstration that enabled visitors to experience the world of the AR/MR Cloud using AR/MR Cloud technologies developed by NTT DOCOMO R&D (**Figure 1**) [1].

(a) Visitors enjoying content on Magic Leap 1, Mirage Solo and iPad.

(b) Actual content displayed on iPad

Figure 1 The DOCOMO Open House 2020 Exhibition

*3 **XR**: A general term for technologies such as VR, AR, and MR that provide new experiences through the fusion of virtual space and real space.

*4 **VR**: Technology that enables users to immerse themselves in a virtual space separate from real space using a head-mounted display or other device.

*5 **AR**: Technology that uses digital technology to add information to real space using a glasses-type terminal or other device.

*6 **MR**: Technology that enables a virtual world to feel more real by more closely integrating real space and virtual space through the use of a glasses-type terminal or other device.

This demonstration showed three types of devices, Magic Leap 1*7, Mirage Solo (video pass-through type VR goggles*8), and iPad, sharing AR/MR content in an exhibition booth that mimics a town. Many visitors wanted to try Magic Leap 1, and many expressions of wonderment were heard as they experienced the world of the AR/MR Cloud with this cutting-edge spatial computing*9 device.

This article describes the technologies required to realize the world of the AR/MR Cloud developed by NTT DOCOMO R&D and introduces future prospects.

2. The Set of Functions Necessary to Realize the AR/MR Cloud

Figure 2 shows the set of functions necessary to realize the AR/MR Cloud. This set of functions consists of three major functions - the self-localization function, the spatial 3D model generation/management function, and the content space management function, which are described below.

2.1 Self-localization Function

Self-localization is technology that recognizes the exact position and orientation of each AR/MR device in real space. Recognizing position makes it possible to superimpose AR/MR content as if it fits into the real space while sharing the position of each device makes it possible to provide interactive AR/MR experiences.

One existing technology that could be used is Global Positioning System (GPS), but the positioning error of GPS is large, especially in indoor environments, and even in an open sky environment with no occlusions (obstructions), GPS has an error of several meters. Therefore, the positioning recognition accuracy of GPS is insufficient to provide

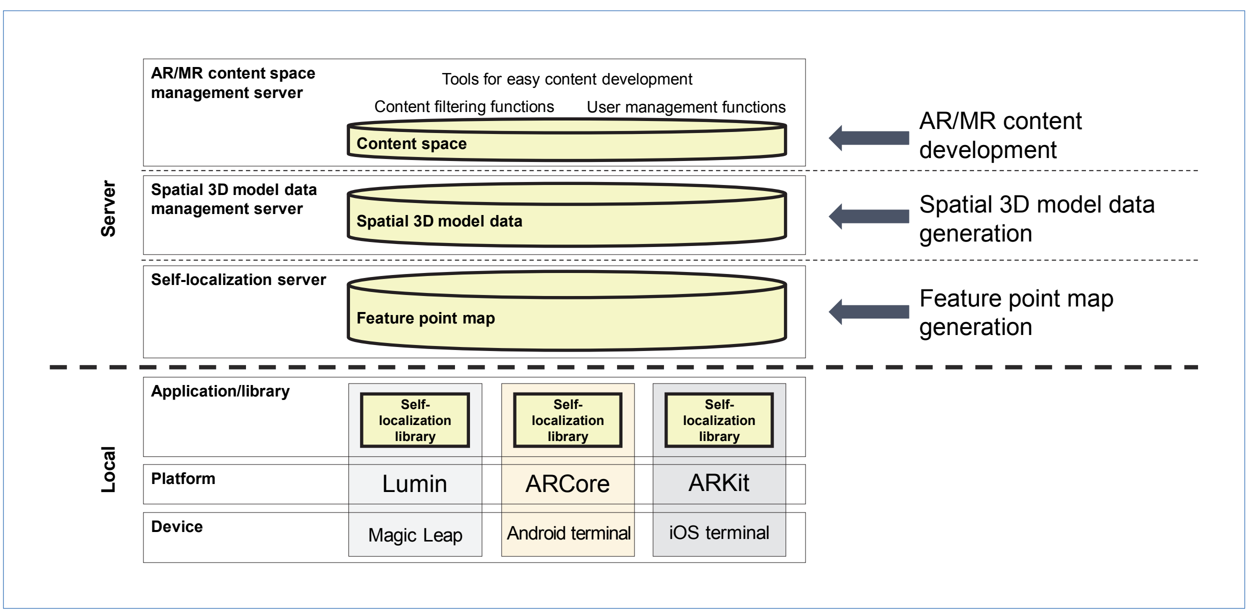


Figure 2　Set of Functions Necessary to Realize the AR/MR Cloud

*7 **Magic Leap 1:** "MAGIC LEAP 1", the Magic Leap logo, and all other trademarks are trademarks of Magic Leap, Inc.

*8 **Video pass-through type VR goggles:** VR goggles that project images recognized by a front-mounted camera onto a display inside the goggles, enabling the wearer to see the outside environment even while wearing the goggles.

*9 **Spatial computing:** Technologies that recognize objects and spaces in the real world and fuses them with digital information. Using these technologies makes it possible to transcend the limitations of two-dimensional displays, integrate real space and the digital world into one, and interact with the digital world in the same way as the real space.

AR/MR experiences in which AR/MR content is accurately superimposed in real space. Another method would be to use AR markers*10 to recognize absolute coordinates in real space, but the issue is AR markers need to be placed so that self-localization can be performed.

To address these issues, NTT DOCOMO has developed a system that enables self-localization by applying Simultaneous Localization And Mapping (SLAM) technology*11, which is a feature point-based positioning technology that uses feature points in images obtained from cameras. The basic process flow of this system is shown in **Figure 3**. The system uses a stereo camera and takes pictures of the area in which to place the AR/MR content. Next, it extracts feature points and creates the feature point map for aligning the AR/MR content with the real world. Then, the system matches the extracted feature points to the coordinates of the real space, makes global coordinates*12, and uploads the feature points matched to the coordinates of the real space (a feature point map) to the server. Meanwhile, each device has a built-in self-localization library, and periodically sends camera images (one every few seconds) to the self-localization server. The server extracts the feature points of the images sent from each device, checks them against the feature point map, and calculates the global coordinates where the images were taken. Using the built-in self-localization library, each device performs real-time tracking of each device's local coordinates and corrects the coordinates in real time using the global coordinates sent by the server. Maintaining a common global coordinate system and

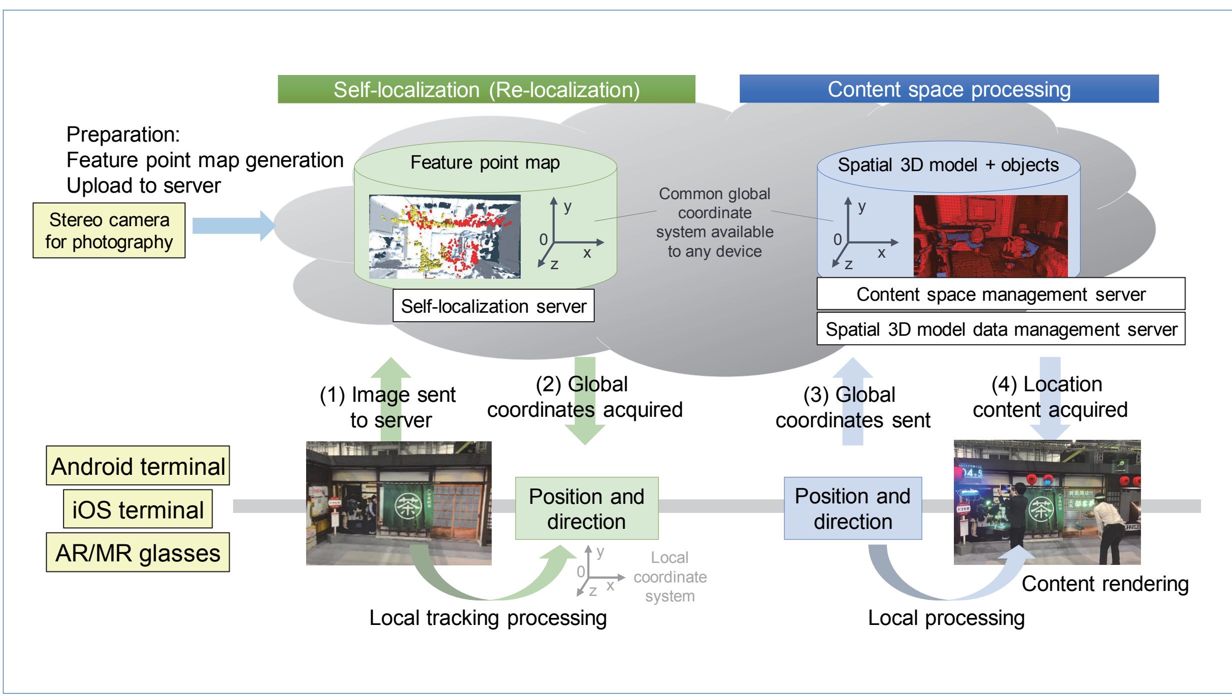


Figure 3　Basic process flow of the system

*10　AR marker: A mark or image used to display digital content on a device screen. Reading AR markers with image recognition technology makes it possible to recognize the location of the AR markers.

*11　SLAM technology: Technology that uses device camera image and sensor information to create a map of a surrounding environment while recognizing the device's own position.

*12　Global coordinates: A coordinate system that represents the entire space where AR/MR content is placed. Transforming the local coordinate system representing the position of each device into a global coordinate system through self-localization makes it possible to view shared AR/MR content across multiple devices.

linking data between the self-localization server, spatial 3D model data management server, and content space management server makes it possible to draw AR/MR content precisely aligned to the location.

In this way, self-localization using SLAM technology enables more accurate position recognition than GPS, without the need to arrange AR markers, etc. Due to the characteristics of SLAM technology, there are some environments where self-localization is easy and others where it is difficult. Thus, performance depends on how the feature point map is created. Currently, the system is capable of self-localization indoors and in some outdoor environments.

In addition, there are issues such as the need to capture images for feature point maps in advance, the photographic know-how required to create highly accurate feature point maps, and the fact that self-localization may not be possible due to changes in the surrounding environment, etc.

2.2 Spatial 3D Model Generation/Management Functions

Spatial 3D model data is used for two main purposes (**Figure 4**). One is to express physical phenomena such as occlusions and bounces of AR/MR content, and the other is to increase visibility for AR/MR content developers for consideration of how and where to place AR/MR content when developing it.

The spatial 3D model data used to represent physical phenomena is not point cloud data, which is a collection of points, but mesh data, which is a collection of surfaces. In the example of throwing a ball in AR/MR content, aligning and arranging transparent surfaces represented by mesh data to match real space makes it possible to express the

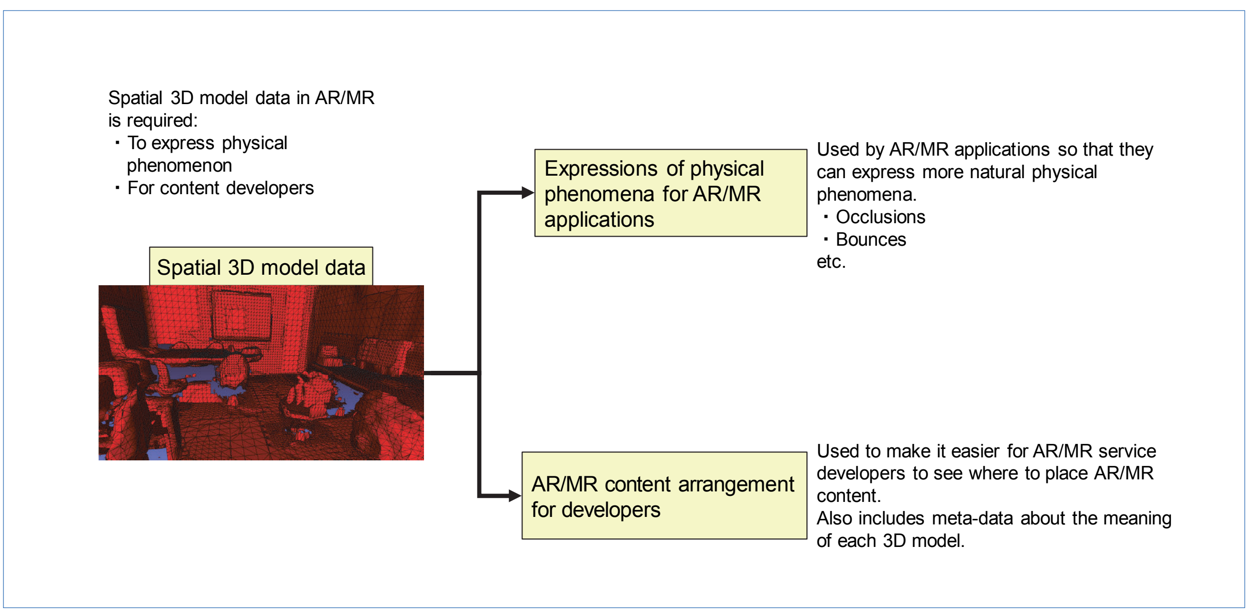


Figure 4　Uses of spatial 3D model data

ball bouncing off walls (bounce) or express rendering (occlusion) such that the ball cannot be seen on the other side of an obstacle. Such expression of physical phenomena is a very important technology to make AR/MR content feel more realistic.

Generating spatial 3D model data by photographing real space in advance and managing it on a spatial 3D model data server makes it possible to share spatial 3D models among users and express physical phenomena through spatial 3D models. As spatial 3D model data generally involves large file sizes, it is important for real-time interaction to ensure the minimum accuracy required to represent physical phenomena and ensure that the file size does not become too large.

Also, developers need to know where cyberspace is in real space so they can place AR/MR content. Since the spatial 3D model data used here only requires a certain level of understanding of locations, point cloud data or mesh data can be used, although high accuracy locations and shapes may be required depending on the content to be placed.

2.3 Content Space Management Function

The content space management function manages the tastes and preferences of device users, as well as attributes data and current status (user management function), and selects and outputs content according to the various attributes and states of the users of each device (content filtering function) (Fig. 2). This function also provides tools to make it easy to develop this content (a user-friendly development environment). The content space in which AR/MR content is placed is managed with a common coordinate system with the feature point map used in self-localization technology and the spatial 3D model data of real space, and has an interface for smooth linking with the self-localization function and the spatial 3D model management function.

3. Conclusion

This article has described the self-localization function, spatial 3D model generation/management function, and the content space management function for realizing the AR/MR Cloud.

NTT DOCOMO R&D is making efforts to develop technologies to realize an AR/MR Cloud world where all users of various devices, including glasses-type AR/MR devices, smartphones, and video pass-through VR goggles, can experience shared AR/MR content. We believe that the world of the AR/MR Cloud - spatial computing where AR/MR content in cyberspace is superimposed onto real space - can provide a completely new and never-before-seen experience. We will continue to develop toward the realization of a world in which all users of AR/MR devices can experience the AR/MR Cloud.

Please go to the official NTT DOCOMO website to find out more about NTT DOCOMO's XR efforts [2].

REFERENCES

[1] M. Tamaoki: "DOCOMO Open House 2020 - Dawn of the 5G era and the Future Beyond -," NTT DOCOMO Technical Journal, Vol.22, No.1, pp.52-60, Jul. 2020.

[2] NTT DOCOMO: "docomo XR|NTT DOCOMO." http://xr.docomo.ne.jp/

Technology Reports R&D Densoku-car Mobile Radio Field Experiment Special Vehicle

History of "Densoku-cars" for Research and Development Supporting the Evolution of Mobile Communications

—Unique Vehicles Used in 1G－5G Mobile Radio Field Experiments—

R&D Strategy Department Yukihiko Okumura

Mobile radio field (outdoor) experiments are an important and indispensable process at the research and development stage of a mobile communication system that aims to provide stable communications under mobile conditions in a variety of environments. For conducting these field experiments, NTT DOCOMO has introduced mobile radio field experiment vehicles (called "Densoku-cars" for short) for research and development purposes having unique and original specifications tailored to each generation from 1G to 5G. Through experiments involving extensive driving over great distances, these Densoku-cars have come to support the evolution of mobile communications.

This article describes the features of Densoku-cars for research and development purposes in each generation and makes extensive use of photos to introduce main Densoku-cars, associated mobile radio field experiments, and auxiliary equipment.

1. Introduction

Radio system engineers in Japan usually call radio wave measurement ("denpa sokutei" in Japanese) or electric field strength measurement ("denkai kyodo sokutei" in Japanese) "densoku" for short in Japanese. Furthermore, vehicles whose purpose is to measure the received state of radio waves for a wide range of areas in a variety of radio systems including mobile communication system and broadcast system are called "Densoku-cars." However, Densoku-cars for research and development (R&D) of mobile communications systems (hereinafter referred to as "R&D Densoku-cars" or "[class name]

Densoku-cars*") described in this article are special vehicles whose purposes are not only to conduct radio wave measurement and electric field strength measurement but also to conduct outdoor mobile radio field experiments such as radio communication/transmission experiments, communication connection tests with commercial equipment, etc. while moving. Their specifications and functions are basically common to those of broadcast relay vehicles used by broadcasters and mobile base station vehicles used by mobile network operators.

Mobile terminals (or mobile stations) used in mobile radio field experiments in the research stage are usually prototype equipment, and in addition to being large in size with high power consumption, they are extremely delicate pieces of equipment. Careful and diverse measures are therefore needed for loading such equipment onto a vehicle. Over the years, NTT DOCOMO's R&D department (including the mobile-communication R&D departments of the Nippon Telegraph and Telephone Public Corporation and NTT eras) have introduced special vehicles to meet the requirements of each generation as R&D Densoku-cars taking the above points into account. The experiences gained from experiments by running Densoku-cars in a variety of environments have driven the evolution of vehicle specifications tailored to the times and have built up extensive know-how in holding field experiments.

Making extensive use of photos, this article describes the features of R&D Densoku-cars, introduces main Densoku-cars and associated mobile radio field experiments that played an active role in each generation of the mobile communication system, and describes auxiliary equipment.

2. Features of R&D Densoku-cars

Photo 1 shows R&D Densoku-cars assembled in the parking lot of the NTT DOCOMO R&D Center (photo taken in October 2007). Against the background of progressively higher frequencies and functionality in each generation of the radio access system, mobile station equipment and measurement equipment for experimental use have become increasingly larger in size. In step with this trend, Densoku-cars have become increasingly larger as reflected by the transition from 3-to-4-ton-class vehicles (small trucks or small buses) to 8-ton-class vehicles (medium trucks) as base vehicles. In addition, measures for making it easier to load large experimental equipment have been taken such as the installation of support rails, movable loading mechanisms, and specialized doors in Densoku-cars. At the same time, the power consumption of experimental equipment has been escalating upward, and to deal with this issue, a large-capacity (maximum 20 kVA*1 class) engine-driven generator*2 (hereinafter referred to as "E/G") that can provide a stable and sufficient amount of power has come to be installed in Densoku-cars. Special measures have also been taken to reduce generator noise and vibration. Furthermore, equipment mounts have been equipped with anti-vibration mechanisms to alleviate the vibration of experimental equipment while the Densoku-car is moving in addition to the adoption of air suspension in the vehicle itself.

Additionally, a variety of measures have been

* Densoku-car class: This article considers vehicle size and scale of equipment to be loaded and divides R&D Densoku-cars into 3 classes (having the class names of small, medium, and large).

*1 VA: Symbol for volt-ampere. This is the unit representing apparent power when supplying an AC power supply to various types of equipment and devices taking into account the active power used by the load and reactive power not used by the load. It is used to rate the capacity of generators, transformers, etc.

*2 Engine-driven generator: An electrical power generator or power generating equipment driven by an engine as a power source that serves to supply power mainly in environments in which commercial power cannot be supplied. Gasoline, gas oil, gas cartridges, etc. may be used as fuel. Called "E/G" for short. When applying generators to R&D Densoku-cars, a variety of methods have been adopted taking into account their use in mobile radio field experiments (see Section 5).

Photo 1　Densoku-cars assembled in the parking lot of the NTT DOCOMO R&D Center

adopted for verification experiments of new radio access technologies and measurement experiments of mobile radio wave propagation characteristics*3 depending on the vehicle. These include the use of detachable specialized mounts that can mount antennas of various shapes according to the experimental items and the use of roof wiring mechanisms for handling antenna cable with low loss. Moreover, to measure and evaluate communication performance and propagation characteristics that depend on antenna ground height, antenna installation position, and other factors, R&D Densoku-cars have been equipped with moveable antenna mounts using extendable poles, folding arms, turntables, etc.

R&D Densoku-cars for conducting special experiments were also introduced. These included Densoku-cars capable of radio communication experiments during high-speed movement up to a maximum of 300 km/h. The idea here was to conduct simulation experiments with the aim of improving communication performance and stabilizing communication quality when using mobile phones or smartphones while riding high-speed trains. They also included Densoku-cars equipped with high extendable antenna poles for conducting radio wave propagation experiments that reproduce the installation conditions of actual base station antennas.

Some R&D Densoku-cars, moreover, have been registered as special purpose vehicles in accordance with the Road Transport Vehicle Act in Japan, that is, as automobiles whose main purpose of use is special in nature. These vehicles are called "number 8 vehicles" because the numbers identifying the type of vehicle on the automobile's license plate (class number) begin with the number 8.

*3　Mobile radio wave propagation characteristics: In mobile communication systems, characteristics such as propagation loss, power delay profile, and angle profile in the propagation of radio waves from a transmission point (base station or mobile terminal) to a receive point (mobile terminal or base station). In land movement, obstacles and reflective objects such as buildings, trees, and undulating terrain are constantly affecting radio wave propagation so that the receive level is always fluctuating according to movement with the possibility of momentary drop offs.

3. R&D Densoku-cars by Mobile Communication System Generation

This section describes the main Densoku-cars that played an active role in R&D in each generation of the mobile communication system.

3.1 1G

On setting out to achieve the first-generation mobile communications system (1G) using a cellular system*4, it was first necessary to clarify and model the mobile radio wave propagation characteristics of the 800 MHz band—a new radio frequency candidate—in the system study. In 1962, researchers at the Electrical Communications Laboratories of Nippon Telegraph and Telephone Public Corporation (Musashino City, Tokyo) clarified the radio wave propagation characteristics between a base station and mobile terminal, the basis of mobile radio link design, through radio wave propagation measurement experiments while actually driving around in an automobile. Then, on the basis of the results obtained, they began studies with the aim of generalizing a method for estimating received electric field strength and service area. Additionally, to obtain actual mobile radio wave propagation data in the Very High Frequency (VHF) band*5 and Ultra High Frequency (UHF) band*6, a large-scale experiment by running Densoku-cars was conducted throughout the Kanto region (which includes the Tokyo metropolitan area) from 1962 to 1965 [1]. The idea here was to conduct mobile radio wave propagation experiments in both urban and suburban areas throughout this region. Transmit equipment was installed at five locations, one of which was the high-reaching Tokyo Tower (Minato ward, Tokyo) in central Tokyo. This location was used as the base point for conducting experiments for propagation distances in the range of 1 – 100 km. The Densoku-car used in these experiments is shown in **Photo 2**. Using the Mitsubishi Jupiter truck as the base vehicle, this vehicle (medium Densoku-car) was equipped with receiving antennas for measuring the reception of

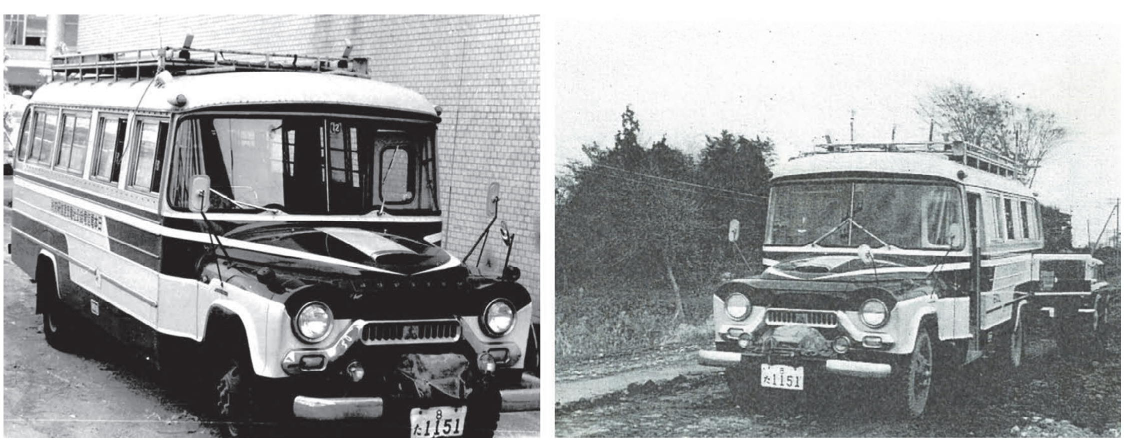


Photo 2 Densoku-car and experimental scene in large-scale radio wave propagation experiments

*4 **Cellular system:** In mobile communications, a system that divides the service area into many small areas called cells and installs a base station in each cell so that users (mobile terminals) can connect with a neighborhood base station and make calls. If a user (mobile terminal) in the middle of a call should cross over into another cell, the current base station will be automatically switched to the base station in the new cell through handover control without interrupting the call. Giving cells a hexagonal shape can make spectrum utilization even more efficient.

*5 **VHF band:** Frequency band in the range of 30 – 300 MHz with wavelengths of 1 – 10 m. Also called the meter-wave or ultra-short-wave band.

*6 **UHF band:** Frequency band in the range of 300 MHz – 3 GHz with wavelengths of 10 cm – 1 m. Also called the decimeter-wave or super-ultra-short-wave band.

radio waves emitted from transmit equipment, received-electric-field-strength measurement equipment, and an E/G for supplying power. The measurement equipment at that time used vacuum tubes in the transmitter and receiver, and if operated continuously throughout the day, measurement values would vary over a range of approximately 10 dB, which required that the equipment be used while performing a level calibration every 30 minutes. The running distance of this experiment conducted twice throughout the Kanto region came to approximately 2,500 km in total.

Next, based on the results of these large-scale mobile radio wave propagation experiments, development of a commercial 800 MHz band land mobile telephone system (metropolitan system) [2] began in 1972 at the Yokosuka Electrical Communication Laboratory (Yokosuka City, Kanagawa Prefecture) of Nippon Telegraph and Telephone Public Corporation. Based on a cellular system, this development work involved the establishment of technologies as a foundation for this system. These included frequency effective utilization technology that adopted Frequency Division Multiple Access (FDMA)*7 in a frequency reuse system*8 and multiple access scheme*9 as well as radio link control technology, mobile telephone switching technology, etc. As part of this development effort, system tests were conducted in the field from 1973 to 1976.

The R&D Densoku-cars used in these tests are shown in **Photo 3**. A variety of tests were conducted using a microbus vehicle (medium Densoku-car, Photo 3 (a)) that could seat seven using the Toyota Massy Dyna (medical examination car model that could carry X-ray equipment; 4-ton vehicle) as the base vehicle and a station-wagon vehicle (small Densoku-car, Photo 3 (b)) with extended space in the rear for equipment using the Toyota Crown Van as the base vehicle. The microbus vehicle was equipped with communications experimental equipment and various types of measurement equipment as well as a computer (equipment installed on the left side in the lower-right vehicle-interior photo of Photo 3 (a)) for analyzing and plotting experimental data. Power was supplied to all of this equipment from a 6 kVA E/G. The station-wagon vehicle, meanwhile, was introduced to reproduce the mounting of a mobile terminal in an ordinary passenger vehicle with the same antenna height and to evaluate the effects of noise when using the car phone (noise in both the high-frequency band and voice band). This type of vehicle was also used in system tests of a small and medium-sized urban system such as for regional cities with a small subscriber base.

On completion of development, the metropolitan system began commercial services in Tokyo’s 23 wards in December 1979 as a car phone service (voice call service), and the small and medium-sized urban system began service in Hiroshima, Gifu, and Sendai in March 1983. Following this, portable shoulder phones and hand-held mobile phones were commercially introduced in 1985 and 1987, respectively. Continuing on, a practical large capacity system [3] that could accommodate five times as many subscribers as the previous system was developed to meet increasing demand with commercial services launched in 1988. Then, in 1991, the super-compact portable telephone “Mova” [4] having a terminal size

---

*7 FDMA: A multiple access scheme in the frequency domain that divides a wide frequency band allocated to a system into multiple narrow band radio channels each of which can be used independently by a different user as a communication channel.

*8 Frequency reuse system: A system for making effective use of limited radio frequency resources in a cellular system by repeating and reusing the frequency used by a certain base station (cell) in another cell located at a certain distance away to avoid inter-cell interference. This scheme divides the entire frequency band allocated to the system into a single group of N frequency bands and repeats those frequency bands across the entire system in such a way that cells using the same frequency band are not adjacent. (As a typical example, $N = 7$ corresponds to 7 reused frequencies.)

(a) Microbus Densoku-car and test scene

(b) Station-wagon Densoku-car and test scene

Photo 3 Densoku-cars used in R&D of 1G (metropolitan system and small and medium-sized urban system)

of only 150 cc was commercially introduced.

3.2 2G

Around the time that 1G was introduced, initial studies on the next-generation system began. In 1987, with the aim of increasing system capacity and improving voice quality beyond 1G and achieving new services such as facsimile and data communications in addition to voice call services, R&D of the second-generation mobile communications system (2G), namely, the Personal Digital Cellular (PDC) system*10 [5] using Time Division Multiple Access

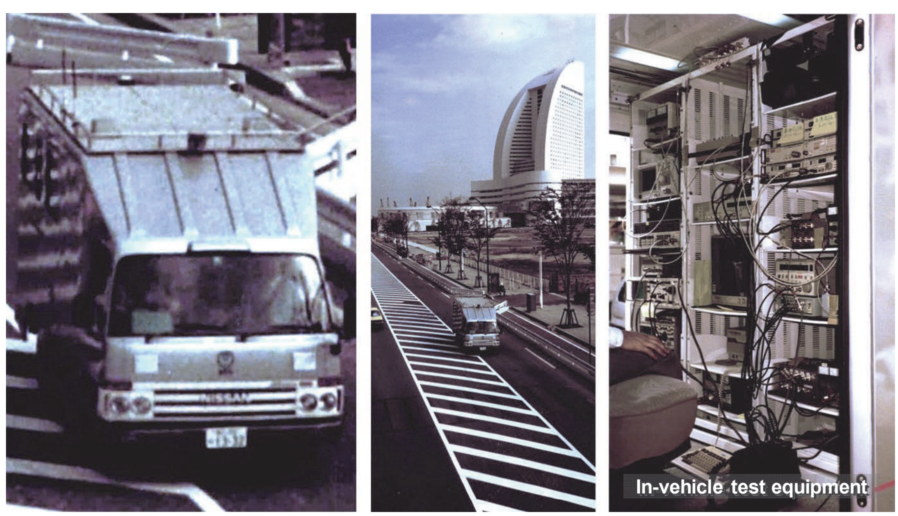
*9 Multiple access scheme: In a mobile communications system, a system that enables a base station to accommodate multiple users (mobile terminals) and that enables each terminal to correctly send/receive information without mutual interference when those terminals are simultaneously performing radio communications with that base station. Methods exist for allocating radio resources to each user by dividing them into time, frequency, and other domains, and various schemes have been adopted in each generation's system to accommodate even more terminals using limited radio frequency resources such as methods that combine multiple domains.

*10 PDC system: A second-generation mobile communications system widely used in Japan adopted by NTT DOCOMO and others.

(TDMA)*11, began in earnest and a variety of mobile radio field experiments were conducted the same as in 1G.

Photo 4 (a) shows a view of an experiment conducted in 1989 in the vicinity of Shinjuku ward, Tokyo on achieving high-quality transmission by combining Quaternary Differential Phase-Shift Keying (QDPSK)*12, a digital transmission system, and selective receive diversity*13. The base vehicle for the R&D Densoku-car used here was the General Motors (GM) Chevy Van (Chevrolet brand) from the United States. This vehicle (small Densoku-car)

(a) Small Densoku-car and QDPSK digital transmission experiment



(b) Medium Densoku-car and system test

Photo 4 Densoku-cars used in R&D of 2G (PDC system)

---

*11 TDMA: A multiple access scheme in the time domain that divides a single radio channel having a prescribed frequency band into multiple fixed time intervals (time slots) each of which can be used independently by a different user as a communication channel.

*12 QDPSK: As a phase-shift keying (phase modulation) system that modulates transmission digital data in the phase domain of the radio signal, a system that uses four phase states per symbol and that reflects data in the phase difference between symbols. In communications via a mobile radio wave propagation path in which amplitude and phase fluctuate greatly due to the fading phenomenon, QDPSK can improve communications quality compared with systems using absolute phase.

incorporated an E/G and a roof-mounted platform for installing antennas equipped with a mechanism for varying the spacing between multiple antenna elements. The reason for using an American vehicle as the base vehicle was that there was no suitable Japanese vehicle at that time that could install prototype equipment for such a communication experiment and an E/G to provide the power supply capacity required for operating that equipment, and that could also drive smoothly considering road conditions in a downtown district.

On approaching the last phase of 2G R&D, a series of mobile radio field experiments were conducted from 1990 to 1992 in Yokohama City, Kanagawa Prefecture near Yokohama Station and in the Minatomirai district. The base vehicle for the main R&D Densoku-car used in these experiments was the Nissan Atlas (4-ton truck) shown in Photo 4 (b). This vehicle (medium Densoku-car) was equipped with a box-shaped measurement room, experimental/measurement equipment racks, E/G, etc. The experimental/measurement equipment racks (Photo 4 (b), right) consisted of three standard racks for electronic equipment forming an integrated unit. The vehicle was equipped with a mechanism for moving this integrated rack unit back and forth on rails embedded in the floor or keeping it fixed depending on the types of equipment accommodated and the target experimental/measurement items.

Commercial services for 2G (PDC system) began in March 1993. Then, in March 1997, commercial services [6] began for the PDC packet system (maximum 28.8 kbps) achieving three times the digital data bit rate of PDC system, and in February 1999, "i-mode" services [7] enabling Web access, mail, etc. began using the same system.

3.3 3G

Research into the third-generation mobile communications system (3G) got under way at about the same time as the completion of 2G development. After comparing and studying multiple candidates, the multiple access scheme chosen for 3G was one based on Code Division Multiple Access (CDMA)*14. It was the wideband coherent Direct Sequence-CDMA (DS-CDMA) [8] featuring wideband radio transmission using the newly allocated 2 GHz band (system employing this scheme, came to be called "Wideband-CDMA (W-CDMA) system*15" later). This system widens the radio frequency bandwidth up to a maximum of 5 – 10 MHz, which makes it easy to separate/combine the paths generated by radio wave propagation in an urban setting (multipath*16 phenomenon) and to achieve high-speed and high quality data transmission. A series of mobile radio field experiments were conducted in Funabashi City, Chiba Prefecture starting in 1994 to check the effect of bandwidth widening in a real environment, evaluate video transmission by maximum bit rates of 384 kbps – 2 Mbps, and test newly adopted technologies such as fast power control and soft handover*17.

Photo 5 (a) shows the R&D Densoku-car used in the initial mobile radio field experiments of the wideband coherent DS-CDMA system featuring the new NTT DOCOMO logo. The base vehicle for this Densoku-car was the Nissan Atlas used in

---

*13 **Diversity:** Technology for suppressing the fading phenomenon (a significant drop off in the received signal level) during terminal movement in a multipath environment by combining or selectively using different fluctuating components obtained by giving the transmit/receive signal redundancy thereby reducing the probability of a drop off in the received signal level and improving radio transmission quality. There are various types of diversity such as space diversity, frequency diversity, and time diversity depending on the domain used to give redundancy.

*14 **CDMA:** A multiple access scheme in the code domain that enables simultaneous communication by multiple users on a radio channel of the same frequency band and time slot by multiplexing a spread spectrum communication channel using different orthogonal code streams for each user.

*15 **W-CDMA system:** The name given to the third-generation mobile communications system (3G) that was made into international standardized specifications at 3GPP.

(a) Medium Densoku-car and wideband coherent DS-CDMA experiment

(b) Newly deployed medium Densoku-car and 384 kbps – 2 Mbps transmission experiment

Photo 5 Densoku-cars used in R&D of 3G (W-CDMA system)

---

*16 Multipath: The phenomenon in which a radio signal emitted from a transmission point arrives at the reception point over multiple paths due to its reflection and diffraction off of obstacles such as buildings and topographical features.

*17 Soft handover: In handover control that automatically switches a user (mobile terminal) in the middle of a call to the base station of a different cell that the user has crossed into, a method of moving between cells with no momentary interruption by simultaneously connecting to multiple base stations in the area near cell borders and combining and receiving the signals from those base stations. In contrast, the method that performs only cell switching without simultaneous connections near cell borders is called "hard handover."

the 2G system tests described above. This vehicle was later replaced by a new type of Densoku-car as shown in Photo 5 (b) for use in transmission experiments up to 2 Mbps considering the need for loading large experimental equipment and conducting long-term experiments by running the Densoku-car. This new vehicle (medium Densoku-car) used the Isuzu Elf (3.5-ton truck) as its base vehicle and came equipped with a measurement room in the vehicle's rear body molded with safety (strength), insulation, and attractive design in mind, a low-noise E/G, and air conditioning equipment as well as air suspension to lessen vibration during vehicle running. This vehicle also added dedicated doors for loading/unloading large experimental equipment and introduced for the first time an equipment-fixing base with rails so that equipment could be slid in from the side of the vehicle using a dedicated lift, a mechanism for directly connecting various types of experimental antennas installed outside the vehicle to experimental equipment, and other auxiliary equipment (described later). This new type of Densoku-car continued to be used or additionally deployed across system generations and the seventh version of this vehicle is currently seeing active use in field experiments.

Commercial services [9] for the 3G (W-CDMA) system began in October 2001 (maximum bit rate at the time of service launch was 384 kbps). This was followed by the launch of commercial services for the High-Speed Downlink Packet Access (HSDPA)*18 system [10] in August 2006 (maximum bit rate on the downlink at the time of service launch was 3.6 Mbps; this was successively increased and the provision of a bit rate of 14 Mbps began in June 2011) and for the High-Speed Uplink Packet Access (HSUPA)*19 system in June 2009 (maximum uplink bit rate of 5.7 Mbps).

3.4 4G

Prototyping of radio access experimental equipment toward the fourth-generation mobile communications system (4G) began in 2002, the year following 3G commercialization [11]. Here, to expand the radio frequency bandwidth to achieve high-speed packet communications in the downlink of 100 Mbps and to eliminate the effects of multipath mutual interference and frequency selective fading*20 that tend to increase due to band widening, NTT DOCOMO adopted Orthogonal Frequency Division Multiple Access (OFDMA) as the base radio access system for 4G. This system divides the downlink signal into multiple narrow-band signals to perform multicarrier transmission*21. Using prototype equipment developed for the purpose of verifying the system, 100 Mbps transmission experiments were conducted in mobile radio field experiments from May 2003.

Photo 6 (a) shows the medium Densoku-car (Isuzu Elf) used in these experiments conducted in Yokosuka City. Using the previously described mechanism for sliding in experimental equipment from the side of the vehicle, these experiments were conducted by reloading multiple types of mobile station experimental equipment.

In May 2005, a real-time packet-signal transmission experiment [12] achieving a maximum downlink bit rate of 1 Gbps with a 100 MHz bandwidth was

*18 **HSDPA**: A high-speed packet transmission system on the downlink based on W-CDMA. The maximum downlink bit rate in the 3GPP standard is approximately 14 Mbps. It optimizes the modulation system and code rate according to the signal reception conditions of the mobile terminal.

*19 **HSUPA**: A high-speed packet transmission system on the uplink based on W-CDMA. The maximum uplink bit rate in the 3GPP standard is 5.7 Mbps. It optimizes code rate, spreading factor, and transmission power according to the signal reception conditions at the base station.

*20 **Frequency selective fading**: Refers to the state of a radio propagation path having multipath characteristics in which the power level on the frequency axis of the received signal is not uniform.

*21 **Multicarrier transmission**: A method for transmitting digital data in parallel using multiple carriers (carrier waves). Orthogonal Frequency Division Multiplexing (OFDM) is a typical example of a system using multicarrier transmission, which has come to be used in many recent radio systems including mobile communication systems from LTE on as well as wireless LAN, digital TV, etc.

(a) Medium Densoku-car and 100 Mbps transmission experiment

(b) Newly deployed large Densoku-car and 1 Gbps transmission experiment

(c) Medium Densoku-car and Super 3G system experiment

Photo 6　Densoku-cars used in R&D toward 4G successfully conducted. The experimental equipment used here applied a system that combined 4 × 4 Multiple-Input-Multiple-Output (MIMO) multiplexing technology*22 and 16 Quadrature Amplitude Modulation (QAM) higher-order modulation technology*23 with the base OFDMA radio access system. Then, in December of the same year, a packet-signal transmission experiment achieving a maximum downlink bit rate of 2.5 Gbps was successfully conducted by extending the radio access system to 6 × 6 MIMO multiplexing and 64QAM, and in December 2006, a subsequent transmission experiment [13] achieving a maximum downlink bit rate of 5 Gbps was successfully conducted by applying 12 × 12 MIMO multiplexing.

Next, to deal with the further increase in equipment size in this series of Gbps-class transmission experiments, a new vehicle (large Densoku-car) based on the Isuzu Forward (8-ton truck) was deployed and used in mobile radio field experiments as shown in Photo 6 (b). This vehicle could install two rows (front-and-rear) of experimental equipment in the measurement room. In terms of vehicles used for communication experiments, this vehicle was the largest yet. It was also equipped with a 20 kVA-class large-capacity E/G and an Uninterruptible Power Supply (UPS) to provide a stable supply of power to radio access equipment performing high-speed and advanced signal processing and to various types of equipment for demonstrating parallel transmission of multiple video streams using high-speed transmission.

In addition to the above, NTT DOCOMO proposed the Super 3G concept [14] to the world in November 2004. The idea behind this concept was to enable a smooth transition to 4G while maintaining a long-term competitive edge in 3G technology. The development of an experimental system for testing Super 3G began in 2006, and after conducting basic indoor experiments, mobile radio field experiments in the vicinity of Yokosuka City and Kofu City, Yamanashi Prefecture began in February 2008 using a medium Densoku-car (Isuzu Elf) as shown in Photo 6 (c). A maximum bit rate of approximately 250 Mbps on the downlink was confirmed by these field experiments. The same experimental system assessed the utility and effectiveness of Super 3G toward commercialization using a variety of applications such as voice/image transmission and games.

Super 3G was standardized as Long Term Evolution (LTE) in 3rd Generation Partnership Project (3GPP) specifications in March 2009 and commercial services in Japan [15] began in December 2010 (maximum bit rate on the downlink at the time of service launch was 75 Mbps; this was successively increased doubling to 150 Mbps by October 2013). Then, in March 2015, the PREMIUM 4G [16] commercial service was launched using LTE-Advanced, a further development of LTE, that applied new high-speed, large-capacity technologies such as carrier aggregation*24 and Advanced Centralized Radio Access Network (C-RAN)*25 (maximum bit rate on the downlink at the time of service launch was 225 Mbps; this was successively increased reaching 1.7 Gbps by March 2020).

3.5 5G

R&D on the fifth-generation mobile communications

*22 **MIMO multiplexing technology:** Signal transmission technology that simultaneously transmits different information signals from multiple antennas using the same frequency and receives the signals by multiple antennas through spatial multiplexing on the radio wave propagation path.

*23 **QAM higher-order modulation technology:** One example of a technology that converts digital data to a radio transmission signal with high efficiency. A 16QAM conversion, which represents the signal after conversion by 16 combinations of different phases and amplitudes, can transmit 4 bits of information in one transmission symbol.

*24 **Carrier aggregation:** A key radio access technology in LTE-Advanced that enables high-speed communication by bundling multiple frequency bands (carriers).

*25 **C-RAN:** Network architecture that centralizes the control functions of multiple base stations.

system (5G) began in 2010, the year of LTE commercialization, with the aim of achieving a system having features such as high-speed and large-capacity transmission, low latency, and massive connectivity. In December 2012, a basic experiment in packet signal transmission achieved a maximum bit rate of approximately 10 Gbps using experimental equipment that increased bandwidth in the 11 GHz band to 400 MHz (four times that of LTE-Advanced) while applying 8 × 16 MIMO multiplexing technology [17].

A mobile radio field experiment conducted in Ishigaki City, Okinawa Prefecture using this experimental equipment showed that the target of ultra-high-speed communication exceeding 10 Gbps by 5G could be achieved. This was accomplished by loading a mobile station (transmission equipment and antennas) onto a vehicle (small Densoku-car) based on the Toyota Estima (hybrid car) as shown in **Photo 7** and conducting broadband radio transmission using high frequencies above 6 GHz that had so far been considered difficult to use considering radio wave propagation characteristics in prior mobile communication systems. In addition, the equipment used in this experiment was supplied power through a configuration that used a 1.5 kVA power supply output from the hybrid-car battery and power from an additionally mounted battery to make up for any deficiencies in power. This configuration made it possible to avoid the noise caused by loading an E/G when driving the Densoku-car through residential streets on the test course.

The full-scale R&D stage for 5G got under way in 2014. At this time, a number of frequency bands were being considered as candidates for 5G frequencies including those from the low Super High Frequency (SHF) band*26 such as 3.7 GHz and 4.5 GHz—which were slightly higher than the current 4G frequencies—to the high SHF band such as 15 GHz and 28 GHz and the even higher Extremely High Frequency (EHF) band*27 such as 39 GHz and 70 GHz. Parallel experiments on radio access technologies in each of these frequency bands were conducted in collaboration with leading vendors from around the world [18].

The low SHF experiments included a transmission field experiment of a distributed antenna system*28 using the 4.5 GHz band and a mobile radio field experiment of a Massive MIMO system*29 in

Photo 7 Small Densoku-car and 10 Gbps transmission experiment

*26 **SHF band:** Frequency band in the range of 3 – 30 GHz with wavelengths of 1-10 cm. Also called the centimeter-wave band.

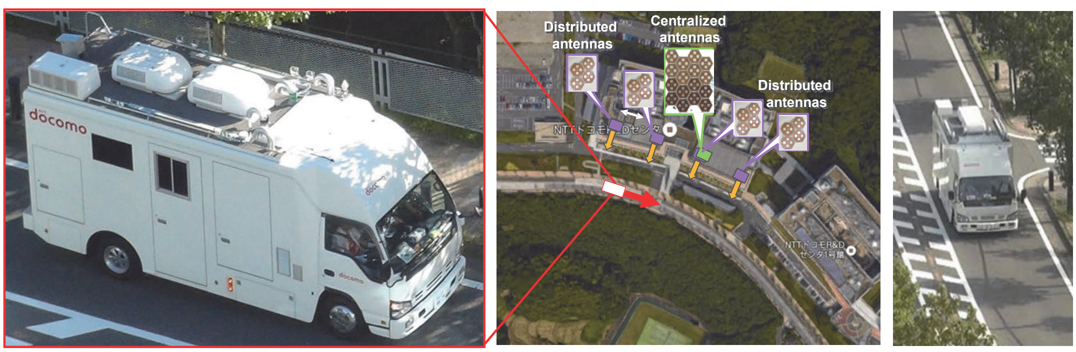
*27 **EHF band:** Frequency band in the range of 30 – 300 GHz with wavelengths of 1-10 mm. Also called the millimeter-wave band (or mmWave as a popular name).

*28 **Distributed antenna system:** A system that can increase communication performance and capacity by connecting multiple geographically distributed antenna units to central-processing (concentrated-processing) equipment via optical fiber and performing integrated signal processing at that equipment.

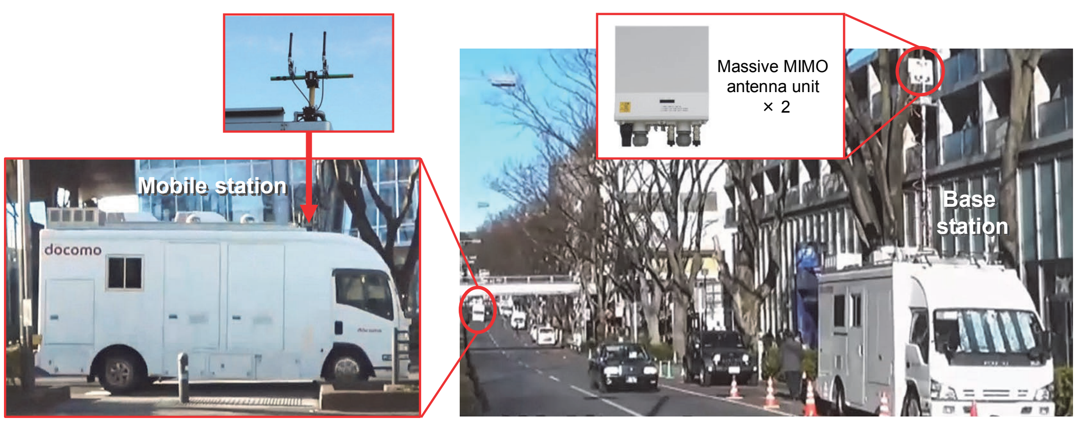
*29 **Massive MIMO system:** A system using large-scale MIMO configured with an ultra-large number of antenna elements. Antenna-element size can be made small in a high-frequency band, so the deployment of this system is being promoted in 5G that uses even higher frequency bands (SHF band – EHF band) than past generations. Corresponds to a centralized antenna system in contrast to a distributed antenna system.

the same 4.5 GHz band with digital beam forming[*30] and beam tracking[*31]. Specifically, in experiments conducted in Yokosuka City and Tokyo in 2016 and 2018, mobile station equipment was loaded onto the medium Densoku-cars (Isuzu Elf) shown in **Photos 8** (a) and 8 (b) to conduct mobile experiments. The beam-tracking field trial was conducted using a medium Densoku-car (Isuzu Elf) parked with permission on the side of the road. This vehicle was equipped with a Massive MIMO base station in which the main unit of the base station was installed inside the vehicle and a Massive MIMO antenna unit[*32] (64 elements × 2) was mounted on an extendable pole.

The high SHF band experiments included high-speed mobile 5G beam-forming communication experiments using the 28 GHz band and a communication experiment using vehicle-mounted 5G glass antennas. First, in an ultra-high-mobility experiment conducted on a racing course (Shizuoka Prefecture) in 2016, the main unit of 28 GHz band mobile station equipment equipped with analog beam forming[*33] was installed inside the small Densoku-car (Toyota Estima) shown in **Photo 9** (a) while the antenna unit



(a) Medium Densoku-car and 4.5 GHz distributed antenna system experiment



(b) Multiple medium Densoku-cars and 4.5 GHz digital beam forming transmission experiment

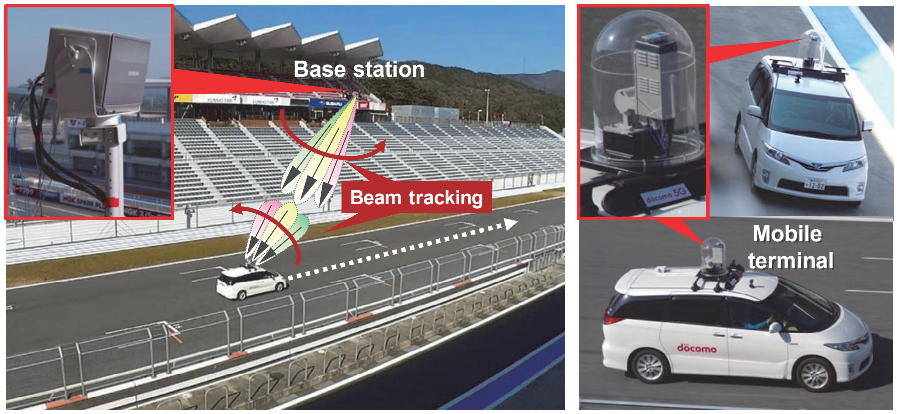
Photo 8　Densoku-cars used in R&D of 5G, (1)

*30　Digital beam forming: A system that uses digital signal processing to achieve a function for concentrating signal radiation in a specific direction (beam forming) using an antenna configured with an ultra-large number of elements.

*31　Beam tracking: A function for following and controlling the radiated direction of a signal subjected to beam forming so that it matches the direction of a moving terminal.

*32　Antenna unit: One piece of equipment for configuring a base station. It integrates the functions for converting the digital signal to be transmitted/received to a radio signal, amplifies the radio signal, and transmits/receives radio signals via the antenna elements.

*33　Analog beam forming: A system that uses analog signal processing to achieve a function for concentrating signal radiation in a specific direction (beam forming) using an antenna configured with an ultra-large number of elements.



(a) Small Densoku-car and 28 GHz band analog-beam-forming ultra-high-mobility experiment

(b) Small Densoku-car and 28 GHz band digital-beam-forming ultra-high-mobility experiment

(c) Small/medium Densoku-cars and 28 GHz band vehicle-mounted glass-antenna communication experiment

Photo 9　Densoku-cars used in R&D of 5G, (2) for the mobile station was installed inside a transparent dome (windproof structure) attached to the roof of the vehicle. In this way, a beam tracking experiment was conducted at speeds of up to 150 km/h envisioning calls while riding high-speed express trains. Next, in an experiment conducted on a test track (Ibaraki Prefecture) in 2020 as part of a millimeter-wave R&D project*34, the mobile station of a 28 GHz band communication-experiment system was installed in the small Densoku-car (Toyota Alphard) shown in Photo 9 (b). This mobile station was equipped with digital beam forming capable of even more accurate beam tracking plus an inter-base station coordination function. In short, an experiment was conducted of a new system that could achieve stable and high-quality communication while overlapping multiple base stations even for automobiles driving on an expressway. In addition, the 28 GHz band communication experiment using vehicle-mounted 5G glass antennas was conducted in an urban area (Sumida Ward, Tokyo) in 2019. In this experiment, the main unit of 28 GHz band mobile station equipment was installed inside the small Densoku-car (Toyota HiAce) shown in Photo 9 (c) while vehicle-mounted 5G glass antennas (two types: glass-integrated antenna*35 and on-glass antenna) were installed on vehicle windows in four places. This mobile radio field experiment conducted at speeds of approximately 30 km/h successfully achieved high-speed radio data transmission at maximum bit rates on the downlink of 3.8 Gbps (400 MHz bandwidth) and 7.5 Gbps (800 MHz bandwidth) [19]. Here, in addition to this mobile station, the main unit of base station equipment was installed inside a medium Densoku-car (Isuzu Elf) parked with permission on the side of the road and a Massive MIMO antenna unit (128 elements × 2 polarization*36 directions) was mounted on an extendable pole (Photo 9 (c)).

In 2017, NTT DOCOMO began its participation in "5G Comprehensive Demonstration Tests (5G Field Trials)*37" of the Ministry of Internal Affairs and Communications (MIC) to test the feasibility of new mobile solutions using 5G in diverse usage fields. This involved the loading of application-oriented equipment such as experimental 5G mobile terminals and high-definition video equipment onto R&D Densoku-cars. For example, in 2018, in the smart workplace field, a test was conducted in Kamiyama Town, Tokushima Prefecture of a "mobile satellite office" (using a small Densoku-car) for performing video editing work while sharing large-capacity content data using 5G [21]. Then, in 2020, in the regional healthcare field, a test was conducted in Wakayama Prefecture of a "mobile health clinic" (using a large Densoku-car) that supports remote medical care while sharing multiple diagnosis video streams [22] (**Photo 10**).

Additionally, from 2019 on, after accelerating the development of mobile terminals and base station equipment toward 5G commercialization and conducting indoor tests, NTT DOCOMO conducted a series of mobile radio field experiments in a variety of communication environments from urban to suburban areas. Then, after the launch of the 5G Pre-commercial Service [23] in September 2019, the 5G commercial service [24] was launched in March 2020 (providing a maximum bit rate on

---

*34 Millimeter-wave R&D project: "R&D for Expansion of Radio Wave Resources (JPJ00024)" conducted by NTT DOCOMO from FY2018 – FY2020 on consignment from the Ministry of Internal Affairs and Communications (MIC).

*35 Glass-integrated antenna: A compact, thin, and transparent glass antenna installed in vehicle windows without obstructing the driver's field of vision or detracting from the vehicle design.

*36 Polarization: Direction of electric-field oscillation when a radio wave propagates through space. A frequently used configuration when transmitting and receiving radio waves from base station antennas consists of both vertical polarization and horizontal polarization in which the electric field vibrates in a plane vertical to the ground and horizontal to the ground, respectively.

the downlink of 3.4 Gbps using the 3.7/4.5 GHz band). After this, an increasing number of models of 5G-compatible mobile terminals began to appear, and in September 2020, services providing a maximum bit rate on the downlink of 4.1 Gbps began using the 28 GHz band for the first time.

Then, in parallel with the spread of the COVID-19 infectious disease from April 2020 on, communication connection tests performed in the field using an R&D Densoku-car as part of the development process for new 5G terminal models were modified so that the collection and analysis of test data (that had previously been performed by a number of measurement personnel within a Densoku-car) could be done from a remote site (**Photo 11**). These tests continue in this form to mitigate the risk of infection.

4. R&D Densoku-cars for Special Experiments

This section introduces R&D Densoku-cars for ultra-high-mobility communication experiments and radio wave propagation experiments as Densoku-cars equipped with the functions needed to conduct special experiments and measurements.

4.1 R&D Densoku-cars for Ultra-high-mobility Communication Experiments

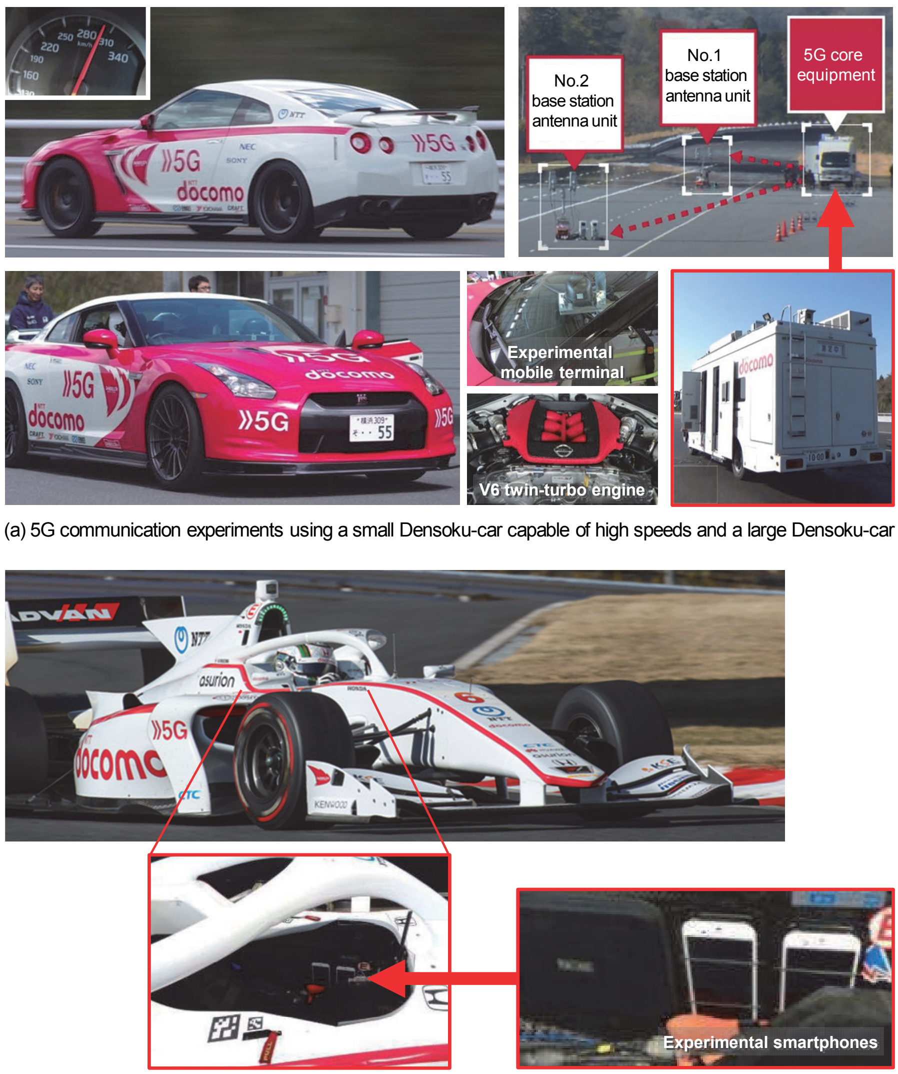
In April 2018, NTT DOCOMO conducted simulation experiments with a running vehicle on the

Photo 10　5G Field Trials by small/large Densoku-cars

Photo 11　Remote testing of 5G commercial mobile terminals by small Densoku-cars

*37　5G Comprehensive Demonstration Tests (5G Field Trials): Tests led by MIC from FY2017 – FY2019 with the participation of mobile network operators and concerned parties in a variety of usage fields with the aim of creating new markets and new services/applications through 5G. The Fifth Generation Mobile Communications Promotion Forum (5GMF) that came to promote and support these tests has published and released reports (in English) on test results [20].

high-speed oval track of the Japan Automobile Research Institute (JARI) in Ibaraki Prefecture to emulate the provision of 5G services in an ultra-high-mobility environment such as high-speed trains moving at speeds in excess of 200 km/h [25]. In these experiments, 5G communications using the 28 GHz band were conducted between experimental mobile terminal equipment loaded on a small Densoku-car capable of running at high speeds and two experimental base stations set up beside the track (**Photo 12** (a)). The base vehicle for this small Densoku-car was a specially tuned NISSAN GT-R sports car



(a) 5G communication experiments using a small Densoku-car capable of high speeds and a large Densoku-car

(b) 5G communication experiments using a formula racing car

Photo 12 Densoku-cars for ultra-high-mobility communication experiment

(mounting a V6 twin-turbo engine) chosen to enable running speeds in excess of 300 km/h while carrying experimental equipment and materials having a total weight of nearly 200 kg. It was equipped with dedicated racks to hold experimental equipment safely and firmly inside the car and mounted a dedicated battery in the trunk to power equipment. These experiments also used a large Densoku-car (Isuzu Forward) to set up quasi-5G core network*38 equipment for controlling handover between the two base stations during communications.

In these experiments, 5G communications were successfully achieved between a base station and mobile station moving at a maximum speed of 305 km/h. In addition, ultra-high-mobility 5G radio data transmission with a downlink bit rate of 1.1 Gbps was successfully achieved while moving at 293 km/h, and handover, that is, switching to another base station while maintaining a radio communication link between the 5G mobile station moving at 290 km/h and a 5G base station, was likewise achieved. A 5G radio live relay of video from a car-window camera was also achieved through 4K High Frame Rate (HFR) video*39 from a 5G mobile station moving at 200 km/h.

NTT DOCOMO has worked to improve the quality of mobile communication services in ultra-high-mobility environments such as high-speed trains in each system generation. To simulate such ultra-high-mobility environments, NTT DOCOMO continues to conduct a variety of communication experiments including handover tests between small mobile terminals (mobile phones and smartphones) mounted on a formula racing car of the type that competes in Super Formula*40 open-wheel car racing (Photo 12 (b)) and peripheral base stations. These communication experiments using a formula racing car are also conducted on a race track, but here, to obtain conditions in which the car running speed, that is, the terminal moving speed, abruptly changes in a short period, the car is first made to accelerate up to 300 km/h on the straight portion of the race course, and then, when entering a curve, to decelerate suddenly to about 50 km/h. In this way, a race track facilitates the efficient collection of test data by circling the same course any number of times. It can also be used to compare and evaluate experimental data such as when adjusting parameters with the aim of improving communication performance.

4.2 R&D Densoku-cars for Radio Wave Propagation Experiments

Photo 13 shows an R&D Densoku-car specialized for conducting radio wave propagation measurement experiments assuming diverse communication environments from urban to suburban. This special Densoku-car (Isuzu Elf) is equipped with two extendable poles with maximum lengths of 14 m and 24 m to enable experimental antennas to be mounted at arbitrary heights above the ground. These maximum lengths are about 2 to 3 times those of antenna-mounting extendable poles mounted on some medium Densoku-cars, which makes it possible to conduct radio wave propagation experiments for a broad range of conditions with respect to base station antenna height. While moving, the 24 m extendable pole can be horizontally

*38 **Core network:** A network comprising switching equipment, subscriber information management equipment, etc. A mobile terminal communicates with the core network via a Radio Access Network (RAN).

*39 **4K HFR video:** High-definition video with a frame rate of 120 fps double that of 4K standard frame rate video. Using a real-time 4K HFR HEVC codec [26], the experiment achieved smooth and high-presence camera video transmission even for a fast-moving scenario in an ultra-high-mobility environment.

*40 **Super Formula:** The top car race in Japan using formula racing cars having a structure in which tires and the cockpit are not covered by the car body. Its official name is Japanese SUPER FORMULA Championship (formerly Japanese Championship Formula Nippon). NTT DOCOMO has been conducting communication experiments since 1999 using the formula racing cars of DOCOMO TEAM DANDELION RACING [27] that competes in this race.

Photo 13 Densoku-car for radio wave propagation experiments mounting high extendable poles

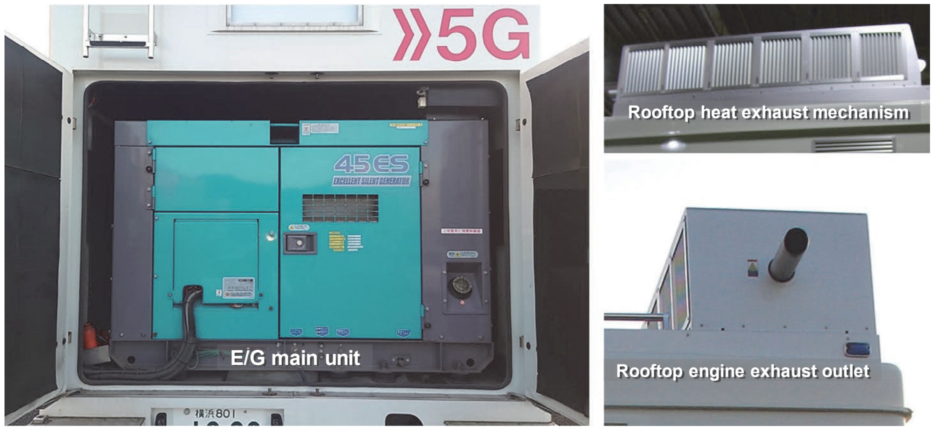
stowed on the roof of the Densoku-car in its shortest contracted state (Photo 13, left). Additionally, when extending an extendable pole near its maximum length, experiments can be conducted while supporting the pole with auxiliary wires at the experiment site.

5. Auxiliary Equipment of R&D Densoku-cars

5.1 Power Supplies for Experimental Equipment

As a major facility in an R&D Densoku-car, the E/G supplies power to various types of experimental and measurement equipment used inside the vehicle. The Densoku-cars of successive system generations (mainly medium and large Densoku-cars) have come to be loaded with E/Gs having a wide range of power supply capacities from several kVA to 20 kVA depending on the power consumption of the equipment and devices used in the experiments. In any case, anti-vibration measures during vehicle travel, soundproofing and anti-heat measures while operating the E/G, and measures for reducing exhaust gas from the E/G were taken reflecting the concern given for tests involving the running of Densoku-cars in urban area, residential districts, etc. (**Photo 14** (a)). Incidentally, for the 1G Densoku-car introduced at the beginning of this article (Photo 2), no domestically made E/G that could be used as a power supply (about 1 kVA) for loaded measurement equipment could be found at that time, so two military surplus products were obtained from the United States military. However, problems in starting up those E/Gs with a starter rope*41 and variations in output voltage made their use difficult, and it was necessary to put aside one unit for repair and inspection while loading the other unit onto the Densoku-car for use in the current experiment. Consequently, if the E/G in use should become

*41 Starter rope: A pull chord used in a recoil starter (manual starter) for manually starting an engine such as an engine-generator. Pulling strongly on this rope rotates the engine's crankshaft and starts the engine. Current engine-generators commonly use a system that starts the engine using a cell motor (see *42).



(a) Low-noise E/G

(b) Power-supply-system central operation panel (top), vehicle-mounted battery (bottom left), UPS and regulated power supply (bottom right)

Photo 14　Power supplies for experimental equipment

unstable, the only way to continue the experiment was to return immediately to Electrical Communications Laboratories and switch that unit with the one whose repair and inspection had been completed. The present E/G is turned over by a cell motor*42 similar to that in vehicle engines and features an automatic voltage regulation function. It is also durable enough for long-term experiments.

Furthermore, as shown in Photo 14 (b), medium and large Densoku-cars loading E/Gs from 3G on are equipped with a panel for centralized operation of the power supply system including start-up of the E/G from within the measurement room. In addition, the introduction of a vehicle-mounted battery system linked to the E/G makes it possible to deal with battery discharging. A large-capacity UPS and stabilizing power supply are also installed for vehicle-mounted equipment. Meanwhile, for small Densoku-cars, the hybrid car*43 is being actively introduced since it mounts a large-capacity battery

*42 Cell motor: A specialized motor for applying torque when starting up an engine. Also called a "starter." The electric power for moving the motor is supplied by a battery but engine startup cannot be performed if the battery is dead.

*43 Hybrid car: A vehicle that has two sources of power: a combustion engine and electric motor. It can improve fuel efficiency by controlling the use of the engine and motor in an optimal manner depending on driving conditions. It is equipped with a dedicated battery for driving the motor that, in some models, can be used as a general-purpose on-vehicle power supply (power supply capacity of about 1.5 kVA). A model that uses a large-capacity battery and enables charging from an external power supply is called a "plug-in hybrid car," which can significantly extend the distance that can be driven with only the motor.

that enables the vehicle itself to supply power to the equipment. By not loading an E/G in this way, NTT DOCOMO is contributing to reduced emission of CO_2, a greenhouse gas. Going forward, the use of electric cars and Plug-in Hybrid vehicles (PHVs) that can reduce emissions even further is desirable.

5.2 Mechanisms for Loading Experimental Equipment

Along with the progress made in R&D in every generation since 3G, the functionality and performance of experimental equipment has dramatically risen and equipment size has become increasingly larger. Newly deployed medium and large Densoku-cars that assume the loading of such large-scale equipment for mobile field experiments have added dedicated doors to the measurement room for loading/unloading experimental equipment and have adopted for the first time an equipment-fixing base with rails so that equipment can be slid in from the side of the vehicle using a dedicated lift (**Photo 15**). The conventional method for mounting various types of experimental equipment in a Densoku-car was to install standard racks for electronic equipment (19-inch rack[*44]) or general-purpose shelves in the measurement room beforehand, place the experimental equipment or measurement equipment on those racks or shelves, and perform the necessary wiring to assemble the experimental system. The new loading style in new Densoku-cars represents a major change over this conventional method. Instead of breaking down integrated experimental equipment in which many circuit cards or modules are incorporated in a large-scale rack (a rack wider than the 19-inch rack or two or three connected 19-inch racks), the new method loads that experimental equipment directly in its existing form. (Subsequent medium and large Densoku-cars have inherited this new style of loading equipment.) Another adopted measure was the insertion of

Photo 15 Equipment-fixing base with rails

*44 19-inch rack: A rack for accommodating various types of electronic equipment with the width of each unit of equipment standardized to 19 inches (482.6 mm) (EIA standard in the U.S. and other standards). The JIS standard in Japan has adopted an equipment width of 480.0 mm but the spacing between holes for equipment mounting is common with the EIA standard (465.0 mm).

vibration-absorption rubber between the equipment-fixing base and the movable platform on the side of the vehicle to counter vibration in prototype experimental equipment requiring careful handling. Air suspension in the vehicle itself was also adopted. These measures make it possible to reduce the effects of vibration having a wide range of frequency components that arise when running the vehicle.

5.3 Mounting Mechanisms for Experimental Antennas

Antennas are a characteristic and important element of experiments targeting radio communication systems using radio waves, and R&D Densoku-cars are equipped with special mechanisms for mounting and using a variety of experimental antennas.

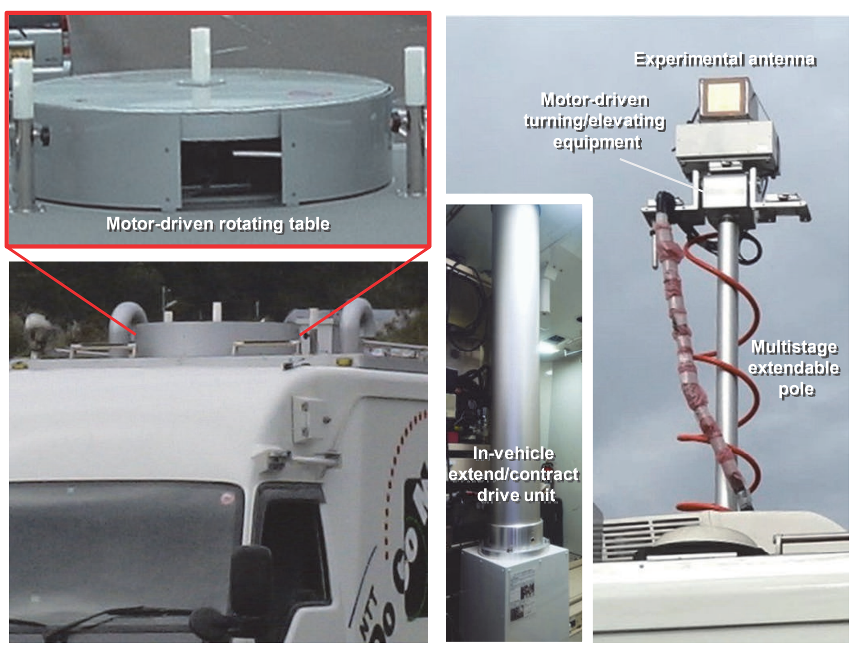
In old-model small Densoku-cars and current medium and large Densoku-cars, experimental antennas have usually been fixed to the deck installed on the roof of the vehicle (roof deck) using specialized fixtures such as magnetic bases or clamps. However, to reduce electromagnetic effects of the metallic vehicle body (including the roof section), some vehicles have been installing antennas using collapsible/projecting pipe bases or fixed projecting rail bases as shown on the left side of **Photo 16** (a) to maintain sufficient separation from the vehicle body.

Additionally, newly deployed medium and large Densoku-cars in 3G and later R&D have come to be used in experiments involving 2 GHz and higher frequencies for the first time given the trend toward higher system frequencies. These vehicles are equipped with roof pass-through pipes (inverted U-shaped pipes) for the wiring of antenna-connection cables as shown at the upper right of Photo 16 (a). These pipes enable direct connection with a relatively short cable length. They negate the need for relay connectors of high-frequency coaxial cable connecting experimental antennas outside the Densoku-car and experimental equipment inside the Densoku-car. They also make for low loss on the cable interval. Here, the reason for using an inverted-U shape is to prevent water penetration by rain or other sources while making the pipe radius large at the pass-through section on the vehicle roof.

On the other hand, small Densoku-cars (excluding some special vehicles and old-model vehicles) are equipped with a general-purpose antenna base as shown at the lower right of Photo 16 (a). For use in mobile radio field experiments, this base adopts a plug & socket attachment that makes it easy to switch antennas of different frequencies for each experiment and to compare and evaluate multiple types of antennas in the same experiment.

There are situations in which antenna installation conditions must be finely adjusted and experimented with depending on the type and content of tests targeting a radio access technology based on a new principle or scheme. Some R&D Densoku-cars are equipped with a variety of antenna-moving mechanisms to meet this need. The photos on the left side of Photo 16 (b) show a low-speed motor-driven rotating table installed on the roof of a medium Densoku-car. This mechanism can move antenna elements in a form that draws a circle on a plane surface and enables antenna interval to be varied in combination with a stationary antenna. The

(a) Various antenna mounting mechanisms



(b) Various antenna moving mechanisms

Photo 16　Mounting mechanisms for experimental antennas photo on the right side of Photo 16 (b) shows a multistage extendable pole with motor-driven turning/elevating equipment (angle adjustment mechanism) that can move and adjust the position of antenna elements or antenna unit in an up/down direction and adjust the direction of radio-wave emission from an antenna (azimuth and elevation).

In the mobile radio field experiments shown in Photos 9 and 10 using high-frequency bands in 5G R&D, antenna units integrated with the high-frequency circuit section came to be installed outside the R&D Densoku-car instead of installing antennas alone and new mechanisms (including a wind-resistant mechanism) for installing such units came to be used.

6. Conclusion

This article introduced with photos the R&D Densoku-cars that have supported the R&D of the mobile communication system from 1G to 5G. At present, NTT DOCOMO is promoting 5G evolution to expand upon initial 5G system technologies and drive R&D toward the sixth-generation mobile communications system (6G). Going forward, we seek to achieve a mobile communication network that can meet user expectations and gain their trust by continuing to operate vehicles of all types and sizes in the field as Densoku-cars in support of these R&D efforts while deploying both existing and upgraded vehicles. From 4G on, NTT DOCOMO has been actively involved in promoting and demonstrating new user applications and services that leverage the improved system capabilities of each generation. In addition, it has recently deployed a new category of R&D vehicles oriented to those new applications and services, but I would like to introduce those vehicles at another opportunity.

Finally, I would like to extend my gratitude to those veterans of Nippon Telegraph and Telephone Public Corporation, NTT, and NTT DOCOMO who graciously provided valuable photos of R&D Densoku-cars for this article including those used in the early R&D of the mobile communication system. In addition, I would like to use this opportunity to express my deep gratitude to all concerned at Shoden Communications Inc., Customized Vehicles Business Department (Kanazawa Ward, Yokohama City, formerly Asuka Electronics Co., Ltd), which has supported the manufacturing and maintenance of R&D Densoku-cars over many years.

REFERENCES

[1] Y. Okumura and E. Omori: "Electric-field Intensity in Terrestrial Mobile Radio and Its Fluctuation Characteristics," The Journal of the Institute of Electronics and Communication Engineers of Japan, Vol.50, No.11, pp.2108–2115, Nov. 1967 (In Japanese).

[2] S. Ito: "Mobile Communications in Japan (Special Issue), Car Phone System," The Journal of the Institute of Electronics and Communication Engineers of Japan, Vol.63, No.2, pp.122–127, Feb. 1980 (In Japanese).

[3] M. Kuramoto, K. Watanabe, M. Eguchi, S. Yuki, and K. Ogawa: "Large-capacity Car Phone System," The Journal of IEICE, Vol.71, No.10, pp.1011–1022, Oct. 1988 (In Japanese).

[4] S. Urabe et al.: Research and Development of Mobile/Portable Equipment, Background to Development of Mova Terminals and Future Outlook," NTT DOCOMO Technical Journal, Vol.2, No.3, pp.6–12, Oct. 1994 (In Japanese).

[5] A. Sasaki et al.: "Special Issue on Digital Mobile Communications System," NTT DOCOMO Technical Journal,

Vol.1, No.1, pp.16–49, Jul. 1993 (In Japanese).
[6] M. Onuki et al.: “Special Issue on Mobile Packet Data Communications System,” NTT DOCOMO Technical Journal, Vol.5, No.2, pp.6–37, Jul. 1997 (In Japanese).
[7] K. Enoki et al.: “Special Issue on i-mode Service,” NTT DOCOMO Technical Journal, Vol.1, No.1 pp.4–30, Oct. 1999.
[8] F. Adachi et al.: “Special Issue on Wideband Coherent DS-CDMA Mobile Radio Access,” NTT DOCOMO Technical Journal, Vol.4, No.3, pp.6–24, Oct. 1996 (In Japanese).
[9] K. Nagata et al.: “Special Article on IMT-2000 Services (1) —Launch of FOMA, the Pioneer of the New Mobile Age—,” NTT DOCOMO Technical Journal, Vol.8, No.3, pp.4–26, Sep. 2001.
[10] Y. Goto et al.: “Special Articles on HSDPA,” NTT DOCOMO Technical Journal, Vol.8, No.3, pp.4–23, Dec. 2006.
[11] T. Masamura et al.: “Special Articles on Fourth-Generation Wireless Access Technology -Toward Next-Generation Mobile Communications Systems-,” NTT DOCOMO Technical Journal, Vol.5, No.2, pp.4–54, Sep. 2003.
[12] K. Higuchi et al.: “Special Articles on 1—Gbit/s Packet Signal Transmission Experiments toward Broadband Packet Radio Access—,” NTT DOCOMO Technical Journal, Vol.9, No.2, pp.25–30, Sep. 2007.
[13] H. Taoka et al.: “Field Experiment on 5-Gbit/s Ultra-high-speed Packet Transmission Using MIMO Multiplexing in Broadband Packet Radio Access,” NTT DOCOMO Technical Journal, Vol.15, No.2, pp.25-30, Jul. 2007.
[14] S. Onoe et al.: “Special Articles on Technology Supporting Large-capacity and High-efficiency Communication in the Flat-rate Era,” NTT DOCOMO Technical Journal, Vol.10, No.2, pp.4–41, Sep. 2008.
[15] M. Kitagawa et al.: “Special Articles on “Xi” (Crossy) LTE Service —Toward Smart Innovation—,” NTT DOCOMO Technical Journal, Vol.13, No.1, pp.4–51, Jun. 2011.
[16] S. Abeta et al.: “Special Articles on PREMIUM 4G —Introduction of LTE-Advanced—,” NTT DOCOMO Technical Journal, Vol.17, No.2, pp.4–30, Oct. 2015.
[17] S. Suyama et al.: “10 Gbps Outdoor Transmission Experiment for Super High Bit Rate Mobile Communications,” NTT DOCOMO Technical Journal, Vol.15, No.4, pp.22–28, Apr. 2014.
[18] Y. Kishiyama et al.: “Special Articles on 5G Technologies toward 2020 Deployment,” NTT DOCOMO Technical Journal, Vol.17, No.4, pp.4–69, Apr. 2016.
[19] M. Inomata et al.: “Vehicle Antenna Technology for Stable 5G Communications without Compromising Vehicle Design—5G Vehicle Glass Antennas—,” NTT DOCOMO Technical Journal, Vol.21, No.3, pp.53–59, Jan. 2020.
[20] Fifth Generation Mobile Communication Promotion Forum (5GMF): “White Paper & Report: 5G System Integrated Verification Trial Report.”
https://5gmf.jp/whitepaper/
[21] Y. Okumura and S. Suyama: “[R&D and Demonstrations toward the Beginning of 5G Era] Co-creation of New Services in 5G Era and 5G System Trials,” IPSJ Magazine, Vol.61, No.3, pp.243-249, Mar. 2020 (In Japanese).
[22] Y. Okumura, J. Mashino, N. Minamida, S. Suyama and T. Okada: “[5G-driven New Communications Society and Supporting Elemental Technologies] Field Trials of Telemedicine System Utilizing 5G,” IEICE Communications Society Magazine (2020 Winter Issue, No.55), Vol.14, No.3, pp.186–199, Dec. 2020 (In Japanese).
[23] E. Endo et al.: “Overview of 5G Pre-commercial Service,” NTT DOCOMO Technical Journal, Vol.21, No.3, pp.4–8, Jan. 2020.
[24] Y. Kojo et al.: “Overview of 5G Commercial Service,” NTT DOCOMO Technical Journal, Vol.22, No.1, pp.4–9, Jul. 2020.
[25] Y. Okumura: “World’s First 5G Wireless Data Transmission in Ultra-high-mobility Environment Exceeding 300km/h,” The Journal of IEICE, Vol.101, No.9, pp.936-937, Sep. 2018 (In Japanese).
[26] NTT Press Release: “NTT Develops World’s First Real–time 4K High Frame Rate HEVC Codec—Enabling live transmission of High Frame Rate video, resulting in smoother and sharper video quality—,” Apr. 2018.
[27] DOCOMO TEAM DANDELION RACING: “DOCOMO TEAM DANDELION Official Website” (In Japanese).
https://www.dandelion-racing.com/team/

NTT DOCOMO
Technical Journal　Vol.22　No.4

Editorship and Publication
NTT DOCOMO Technical Journal is a quarterly journal edited by NTT DOCOMO, INC. and published by The Telecommunications Association.

Editorial Correspondence
NTT DOCOMO Technical Journal Editorial Office
R&D Strategy Department
NTT DOCOMO, INC.
Sanno Park Tower
2-11-1, Nagata-cho, Chiyoda-ku, Tokyo 100-6150, Japan
e-mail: dtj@nttdocomo.com