

International Workshop on Emerging ICT (IWEICT 2023) Program Book



Future Hall (the workshop venue)

Foreword

It is my great pleasure and honor to welcome you to the 20th International Workshop on Emerging ICT (IWEICT 2023), held on October 16-18, 2023, at Tohoku University, Sendai, Japan. IWEICT is a trilateral workshop in which Tohoku University, Yeungnam University, and Dalian University of Technology have participated. This workshop aims to provide a platform for graduate and undergraduate students from the three universities to exchange ideas, share experiences, and foster collaborations in information and communication technology (ICT).

ICT is one of the most dynamic and influential fields in modern society, which has brought tremendous benefits and challenges to various domains such as education, health, industry, environment, and culture. The rapid development of ICT also poses new opportunities and problems that require innovative solutions and interdisciplinary approaches. Therefore, this workshop covers a wide range of ICT-related topics, such as artificial intelligence, signal processing, computer vision, wireless and optical communication, Internet of Things, cybersecurity, multimedia, human-computer interaction, and social computing. This year's workshop features three keynote speeches by distinguished professors, 12 oral presentations, and 12 poster presentations by the students of three universities.

I would like to express my sincere gratitude to all the participants for their contributions and enthusiasm. I would also like to thank the organizing committee, the program committee, and the staff for their hard work and support. I hope you enjoy the workshop and have a fruitful and memorable experience in Sendai.

31 October, 2023

Akinori Ito

Chair of IWEICT 2023 Professor of Information Sciences Tohoku University

A handwritten signature in black ink, reading "Akinori Ito". The signature is written in a cursive, flowing style.

Workshop Schedule

31 October (Tue)

8:30	Registration at Future Hall	
9:00-9:10	Opening address Prof. Akinori Ito (Tohoku University)	
9:10-9:45	Keynote Speech 1 Prof. Shuichi Sakamoto (Tohoku University) Virtual acoustic space synthesis using spherical microphone array	Session chair: Prof. Wooguil Pak (Yeungnam University)
9:45-10:20	Keynote Speech 2 Prof. Zhenrong Du (Dalian University of Technology) Remote sensing in land dynamic monitoring	
10:20-10:35	Coffee break	
10:35-12:05	Oral Session 1	Session chair: Prof. Zhenrong Du (Dalian University of Technology)
12:05-13:30	Lunch Break	
13:30-14:00	Poster Session 1	Session chair
14:00-14:30	Poster Session 2	
14:30-15:00	Coffee break	
15:00-15:35	Keynote Speech 3 *Online Prof. Hyun-Chul Choi (Yeungnam University) Applications of Deep Neural Network, Computational Imaging and Semiconductor Manufacturing	Session chair: Prof. Shinichiro Omachi (Tohoku University)
15:35-17:20	Oral Session 2	Session chair: Prof. Takashi Nose (Tohoku University)
17:20-17:30	Closing remark	
18:30-20:30	Banquet (ANA Holiday Inn)	

Workshop Program

Keynote Speech 1 (9:10~9:45)

Virtual acoustic space synthesis using spherical microphone array

Prof. Shuichi Sakamoto (Tohoku University)

Keynote Speech 2 (9:45~10:20)

Remote sensing in land dynamic monitoring

Prof. Zhenrong Du (Dalian University of Technology)

Keynote Speech 3 (15:00~15:35)

Applications of Deep Neural Network, Computational Imaging and Semiconductor Manufacturing

Prof. Hyun-Chul Choi (Yeungnam University)

*Online

Oral Session 1 (10:35~12:05)

10:35~10:50

A preliminary study on photo-realistic facial animation generation using keypoint features

Zikai Shu, Takashi Nose, Akinori Ito (Tohoku University)

Recently, with the increasing demand for automatic synthesis of animated images from text and sound, photo-realistic facial animation generation has become a hot research field. However, the existing methods use OpenFace, which can only be applied to facial data, and there is a problem that when we use whole body data, the face is too small to recognize and label landmarks; Meanwhile, even if we use upper body data, there is still a problem of poor generation of other parts of the image except for the face. In this article, we introduce a new deep learning framework for image animation with key point features. This framework can use landmarks as full body images when we use full body images, and when we use upper body images, besides the face, other parts can also be annotated with landmarks to guide the learning and training process. At the same time, this article introduces a new network structure to help us generate realistic animated images. As a preliminary study on using keypoint features for animation image synthesis, we attempt to use code to extract keypoints from the whole body image using Gaussian heat maps as keypoints, in order to verify its ability to extract features from the whole body image. Afterwards, we used data from the upper body portrait and performed keypoint extraction using the same method to observe its effectiveness.

10:50~11:05

Design of four-elements patch array antenna

Yusuke Ishikawa, Keisuke Konno and Qiang Chen (Tohoku University)

A planar array antenna is an antenna with multiple radiating elements arranged in a specific plane, and the desired radiation pattern can be obtained by controlling the amplitude and phase of the excitation voltage of each element. In this report, we design four-elements planar array antennas and their radiation patterns are demonstrated.

11:05 ~ 11:20

Resilient Power Supply System for B5G Distributed Antenna Systems

Qiongyan Tang, Hirohito Yamada (Tohoku University)

We proposed and experimentally demonstrated a novel power supply system for distributed antennas of the next-generation mobile communication systems. The power supply system is capable of operating distributed antennas with more than 50% solar-generated power and can operate the antennas for at least 3 days even when losing power supply from utility grids.

11:20 ~ 11:35

Meta Transfer Learning used in Low Resource Speech Recognition

Rui Zhou, Akinori Ito and Takashi Nose (Tohoku University)

Recent years have witnessed a growing interest in low-resource speech recognition to develop robust and high-precision models using limited data. However, due to the scarcity of data, achieving high levels of accuracy continues to pose significant challenges. This paper proposes a novel approach to improve automatic speech recognition (ASR) performance in low-resource languages by leveraging meta-transfer learning, which integrates the advantages of transfer learning and meta-learning to ameliorate the training bottlenecks and overfitting issues associated with pre-training models. We conducted extensive experiments on the Common Voice Corpus and Globalphone Corpus to compare the performance of meta-transfer learning with meta learning and transfer learning. The results show that for near-zero resource experiments, our method decreased CER by 7.9% compared to meta-learning, and by 4.6% compared to transfer learning, demonstrating the effectiveness of our meta-transfer learning.

11:35 ~ 11:50

Evaluation of Environmental Sound Classification using Vision Transformer

Changlong Wang, Akinori Ito, Takashi Nose (Tohoku University)

Recently, attention-based vision transformers have reached a huge success on audio classification tasks. Most of the well-performed ones use elaborated model settings and training tricks to get a satisfactory score on a dataset, while the practicability of vision transformers with respect to model settings and training cost are rarely discussed and lack experimental proofs.

We have made a comparative study that boosting vision transformers on ESC-50 step by step, by which to provide common practitioners a solid basis to adapt vision transformers to the general Environmental Sound Classification task. Specifically, model setting, data augmentation, cross-field transfer learning and model pruning are included in our research, and comparative experiments in different parts by steps have been accessed as a solid basis in practice.

11:50 ~ 12:05

A Study of a Method for Improving Cartoon Drawing Using Stable Diffusion

Rei Takahashi, Takashi Nose, Akinori Ito (Tohoku University)

With the spread of image and text generation AI services such as chatGPT, Midjourney, and StableDiffusion, there have been attempts to use AI for contents such as comic books and illustrations. In this study, we investigate a method to

improve the quality of comic books using AI technology based on StableDiffusion. This method aims to enhance every element of a manga work by using AI while preserving the authorship and style of the work, thereby improving the efficiency and quality of the work. The method proposed in this presentation is intended to take a rough illustration and clean it up to match the style of a particular manga.

Poster Session1 (13:30 ~ 14:00)

3D turn-based strategy game using Unity

Wondam Kim, Yunho Kim (Yeungnam University)

This paper presents the development of a 3D turn-based strategy game (TBSG) that emphasizes randomness, strategy, and immersion. The game was developed using the Unity game engine, and the development process was divided into two phases: design and development. The design phase involved designing the behavior of all components that users would interact with in the game, the direction to guide users, quests that guide users according to the planner's intentions, balance design of users, monsters, and items, and level design that constitutes the difficulty. In addition, the process of creating a data table and transferring the data table to a class in C# scripts and a table in a database was described. The development phase involved implementing the desired behavior of the UI, camera controls, character controls, map generation, data processing that needs to be processed during battle, data processing when creating the screen to be displayed to the user, and audio and UI that are played when a specific button or area is selected.

In addition, the paper also discusses the design intention of choosing the Unity editor version and the relationship between classes so that these functions do not affect the program performance. The development process described in the paper explored the importance of game development and various technical challenges through the development of a 3D turn-based strategy game that takes into account randomness, strategy, and immersion. As a result, the possibility in the field of game development was expanded, and the importance of the Unity engine, planning, and development process was emphasized.

Image scanning cashier

Daejun Lee (Yeungnam University)

The barcode or QR code of the commercial product is used at the checkout counter, but the location of the code varies for each product. It is inconvenient to find the location of the code every time. The goal of the Image scanning cashier is to quickly and accurately recognize product information using the image of the product without using the code at the checkout counter to shorten the waiting time for customers. It is more economical to use an embedded board than to use a PC at the checkout counter. Since the computing power of the embedded board is small, there are restrictions that the computational cost of the classifier backbone is small, but it must produce reliable performance. ResNet-50 was used as a classifier backbone and images obtained using webcam were used as inputs. The object of class "Manager" was used to determine the type of product purchased using the outputs of the classifier backbone.

Survey of Digital Image Processing

HUO Wenxiang (Dalian University of Technology)

Digital image processing, which refers to the usage of algorithms to process digital image, has attracted the interests of many researchers in last decades. Common digital image processing includes image enhancement, restoration, encoding, and compression. It is widely used in information transmission, facial recognition, medical and aviation fields. This survey focuses on some important applications of digital image processing and introduce the simple principles behind them. Also,

the developing algorithms and applications of digital image processing with artificial intelligence is introduced. Finally, the prospect of digital image processing is concluded.

The Design and Implement of An AM Radio

Xin Li (Dalian University of Technology)

A radio is a machine which receives amplitude modulation, frequency modulation and other types of radio signals sent by radio broadcasts and extracting these signals from electromagnetic waves, then converting them into sounds that humans can hear. According to the different types of radio broadcast, radio also could be divided to different categories, FM and AM for instance. AM, as the abbreviation of amplitude-modulated, is supposed to transmit audio signal by adjusting the amplitude of the RF-Radio frequency signal. Compared with FM radio, the advantage of AM radio is the easy detection of the signals and the narrow bandwidth. This paper describes the design and implement of an AM radio, which involves the framework, circuitry basics, design principles, soldering and radio debugging.

Water Purifier Capacity Recommendation System Using MobileNetV3

Min Jung Kim, Ji Woo Shong and Gyu Sang Choi (Yeungnam University)

Recent advancements in deep learning technology have brought about significant changes in the household electronics sector. However, there is a lack of appropriate safety services for home water purifiers, which can pose a risk of burns, especially to children and the elderly. According to surveys conducted by the Fair Trade Commission and the Korea Consumer Agency, home water purifiers rank second in terms of the highest occurrence of safety accidents involving children among household appliances. In response, this study proposes a system that recommends the water purifier's capacity to ensure the safety and convenience of use, particularly for children and the elderly. To recommend the capacity, a deep learning model for object detection is required, demanding significant computational resources and memory for high performance. However, embedding a relatively lightweight model such as MobileNetV3 into the water purifier can reduce the model's size and power consumption. Therefore, this system aims to provide users with convenience and tailored services, while minimizing the risk of burn accidents by integrating various embedded modules into a high-quality service.

Two factor authentication system based on transmission of separated images

TaeHui LEE and Seung Yeob NAM (Yeungnam University)

With the increasing prevalence of online information storage and the myriad of activities on the internet, there is a growing demand for enhanced security in the storage and preservation of critical information, both for individuals and businesses. In response to this demand, this paper introduces login methods previously non-existent, marking a new frontier in the realm of login security. Additionally, It introduces a new encryption method that combines the Vigenere and Cipher Block Chaining (CBC) techniques.

Poster Session 2 (14:00 ~ 14:30)

A Survey on Audio Detection Technology

Chen Hao (Dalian University of Technology)

Audio detection technology aims to automatically identify and analyze sound events, representing a crucial branch of digital signal processing. It has extensive applications such as audio emotion analysis, deepfake audio detection, and audio content recognition. Researches in audio detection technology mainly focus on three aspects, which are audio feature extraction,

fundamental algorithms for audio analysis, and the audio data resource constructions. This survey discusses and summarizes the audio detection algorithms in detail. Finally, the prospect of audio detection is concludes.

A Survey on Affective Computing

Xiaoyan Hu (Dalian University of Technology)

Affective computing makes computers have higher and comprehensive intelligence, which purpose is to establish a harmonious human-machine environment. Affective computing give computers the ability to recognize, understand, and express human emotions through the development of technologies such as psychology, wearable technology and deep learning. The research of affective computing mainly consists of the collection of affective databases, the analysis and recognition of affective information and fusion algorithm of emotion signals. This survey focuses on the above three aspects and discusses five different approaches of emotion recognition techniques in detail. Further discussion falls into the unsophisticated emotion recognition source of the current affective computing because of the rough database. Finally, a prospect of affective computing related to fin-grained emotion classification, comprehensive datasets construction and multimodal fusion recognition is given.

Survey on Image Classification

Zhao Yihang (Dalian University of Technology)

Image classification is one of the most important tasks in the field of computer vision, with the goal of dividing input images into different predefined categories. With the rapid development of artificial intelligence technology and the widespread application of deep learning models, image classification plays an important role in many fields such as medical image analysis, face recognition, and image retrieval. This survey focuses on investigation of image classification technologies based on widely-used machine learning methods, including support vector machines (SVM), random forest, and deep learning technologies, convolutional neural networks (CNN) for instance. Firstly, the basic concepts and processes of image classification were introduced, and then the functions and characteristics of various technologies were discussed in detail. The performance of different algorithms is compared in detail. Finally, various challenges and future prospects related to data imbalance, few-shot scenarios and fine-tune large model in image classification is discussed. Future research can further explore cutting-edge issues such as data imbalance and small sample image classification. Overall, this review paper provides useful guidance and reference for further improving the accuracy and efficiency of image classification.

Unmanned car rental service YU-RentCar

Choi YoHa, Seol WooHyeok, Kwon YeongJun, Kang ChangHan, Ho DaeHwan (Yeungnam University)

With the relaxation of COVID-19, outdoor activities, including travel, have seen an increase, leading to growth in the transportation and mobility market. As a result, the car-sharing market has also experienced growth. Another market that has seen growth in the wake of the pandemic is the contactless and automation market. Automation systems like kiosks have been introduced in restaurants and stores to enhance customer convenience, efficiency, and accuracy. So, we have conceived a business idea that targets both of these growing markets that automated car rental. This business allows customers to search and reserve vehicles through a website and kiosks. For administrators, a dedicated web platform is provided for streamlined business and parking management. Through this approach, both customers and administrators can experience the benefits of automation in a car-sharing business.

A Study on the Effectiveness of Different Recognizers in Low Resource Speech Recognition Using wav2vec2.

Takaki Koshikawa, Akinori Ito, Takashi Nose (Tohoku University)

wav2vec2.0 is pre-trained on large amounts of unlabeled data by self-supervision learning, and shows great potential for low-resource speech recognition by its powerful expressive capabilities. In this study, we perform fine tuning of the wav2vec2.0-base model pre-trained on a 960-h Librispeech corpus using a dataset of 8 languages, about 1h each, and investigated the recognition performance of speech recognition for each language when the recognizer is changed to an RNN or LSTM. As a result, the mean value of CER was lower when using a single fully connected layer or biLSTM, but the recognizer that CER is lowest depends on the language.

Development of a Data-Driven Web Platform for Novice Investors: Enabling Informed Stock Investment Decisions

Jihun Son and Cheongwoo Seo (Yeungnam University)

This paper introduces the development of a data-driven web platform designed for novice investors, enabling them to make informed stock investment decisions. Leveraging KRX, DART, and Kiwoom Securities APIs, it collects vital data for publicly-listed companies in KOSPI and KOSDAQ, including stock prices and financial statements. The system calculates stock prices appropriate for corporate value based on the S-RIM model and presents this information on a React-based web platform. This project combines data collection, financial analysis, and web development to offer a user-friendly solution tailored to the needs of new investors, allowing them to access investment reference data easily, empowering them to make informed investment choices.

Oral Session 2 (15:35 ~ 17:20)

15:35 ~ 15:50

Language-Driven Open-World 3D Scene Understanding with Semantic-Level Supervision

Juan Wang, Zhijie Wang, Tomo Miyazaki, Shinichiro Omachi (Tohoku University)

Open-world semantic-level scene segmentation aims to locate and recognize unseen objects, i.e., categories not labeled in the training data. This task is challenging since the model needs to be generalized enough to recognize unseen classes. Despite recent breakthroughs in 2D open-world perception tasks, their straightforward transfer to the 3D domain has been insufficient because of the difference in data amounts. There are massive 2D image-text pairs, but not as many in the 3D environment because of the great labeling effort. To address this challenge, we explore leveraging current 2D open-world segmentation models to enhance the generalization capabilities and open-vocabulary scene understanding of 3D models without additional manual 3D labeling. More specifically, we utilize existing, pre-trained large 2D semantic segmentation models to generate pixel-wise labels of multi-view images and project them into the 3D point clouds for pseudo-label generation based on the camera matrix and depth information. Our method can establish explicit associations between 3D points and 2D images, thus utilizing the enriched category information in images to provide our models with fine-grained supervised signals for open-world recognition.

15:50 ~ 16:05

Robust Federated Learning Based on Multiple Anomaly Model Filters

Wang Bo, Dai Xiaorui, Yang Zi, Wang Zhaoning, Zhang Maozhen (Dalian University of Technology)

Federated learning is a distributed machine learning paradigm that allows multiple participants to collaboratively train models without concerns about data leakage. However, many studies have shown that federated learning is susceptible to attacks by malicious clients. While many existing defenses perform well in the i.i.d datasets, federated learning data may

come from different distributions (non-i.i.d), which affects the performance of existing defense methods. Therefore, we propose a robust aggregation algorithm, AMF-CFL, designed specifically for federated learning in the non-i.i.d setting. The algorithm filters out potentially anomalous behavior in model updates before aggregation. It uses a two-step filtering process to identify benign update values through multi-k-means filtering and using z-score observation. We evaluate its performance against four types of untargeted attacks and two types of targeted attacks. It effectively safeguards the integrity of the global model.

16:05 ~ 16:20

AN ADAPTIVE ALGORITHM FOR TRACKING THIRD-ORDER COUPLED CANONICAL POLYADIC DECOMPOSITION

Xin-Tong Liu, Xiao-Feng Gong, Dong Zhao, Qiu-Hua Lin (Dalian University of Technology)

Coupled canonical polyadic decomposition (C-CPD) of multiple tensors is a fundamental tool for multi-set data fusion. Existing C-CPD works are mainly limited to batch processing techniques for stationary models, yet in practice the C-CPD model may be dynamic and thus adaptive C-CPD tracking techniques are in urgent need. In this paper, we consider the problem of adaptive tracking of a time-varying third-order C-CPD model, and propose a recursive least squares (RLS) based adaptive tracking algorithm. Theoretical and experimental results are provided to show the merits of the proposed C-CPD tracking algorithm over batch C-CPD algorithm and CPD tracking algorithm, in terms of improved accuracy, reduced complexity, and more relaxed working conditions.

16:20 ~ 16:35

Target Localization based on Multistatic MIMO Radar via Double Coupled Canonical Polyadic Decomposition

Guo-Zhao Liao, Xiao-Feng Gong and Qiu-Hua Lin (Dalian University of Technology)

This paper considers target localization with a multistatic MIMO radar system of multiple transmit arrays and multiple receive arrays. We formulate the matched filtered output data into tensors that admit the double coupled canonical polyadic decomposition (DC-CPD) model, which efficiently characterizes the coupling between receive arrays and that between transmit arrays, and multilinear structure in each tensor. We propose a novel algebraic DC-CPD algorithm based on coupled rank-1 detection mapping, and provides analysis into the uniqueness conditions. A post-processing approach is also introduced to calculate and fuse the DOD and DOA information from the DC-CPD results to finally obtain the target locations. Simulations are provided to illustrate the merits of proposed method over CPD and C-CPD methods, with regards to accuracy and identifiability.

16:35 ~ 16:50

Analysis and Data Augmentation for Conversational Speech Synthesis Based on Dialogue Situation

Shoki Kawanishi*1, Yuya Chiba*2, Akinori Ito*1 (*1 Tohoku University, *1 NTT)

Speech synthesized by general text-to-speech systems sounds somewhat unnatural when listened to as dialogue speech. This is due to the system's inability to produce speech that matches the dialogue situation and the fact that the speech synthesis model is learned from monologue speech. To address these issues, this paper investigates the characteristics of speech in different situations in actual dialogue data, and proposes a method to generate dialogue speech of specific speakers by voice conversion, which is generally not easy to collect in large quantities, and use it for dialogue speech synthesis. Experimental results suggest that different speakers may speak in different ways in different situations and confirm that the proposed method can synthesize dialogue-like speech.

16:50 ~ 17:05

DDGAN: Driving dataset generation by Generative Adversarial Network

KyeongMin Park, GyuSang Choi (Yeungnam University)

Recently, autonomous driving technology has been becoming increasingly integrated into our real-world lives. Many automotive manufacturers are applying driver assistance technologies to a variety of products, and regulations are also pushing for the mandatory implementation of driver assistance safety features. However, there remains a significant challenge in autonomous driving technology research, which is the limited availability of datasets and the high cost associated with data collection. To address this issue, our study explores the enhancement of Generative Adversarial Network (GAN) techniques for generating images based on existing driving data. We propose a new neural network architecture, building upon DriveGAN, to generate road environments and surrounding objects more accurately, and we provide a detailed analysis of the results. Through this research, we aim to offer a solution to the critical problem of data scarcity in autonomous driving technology research.

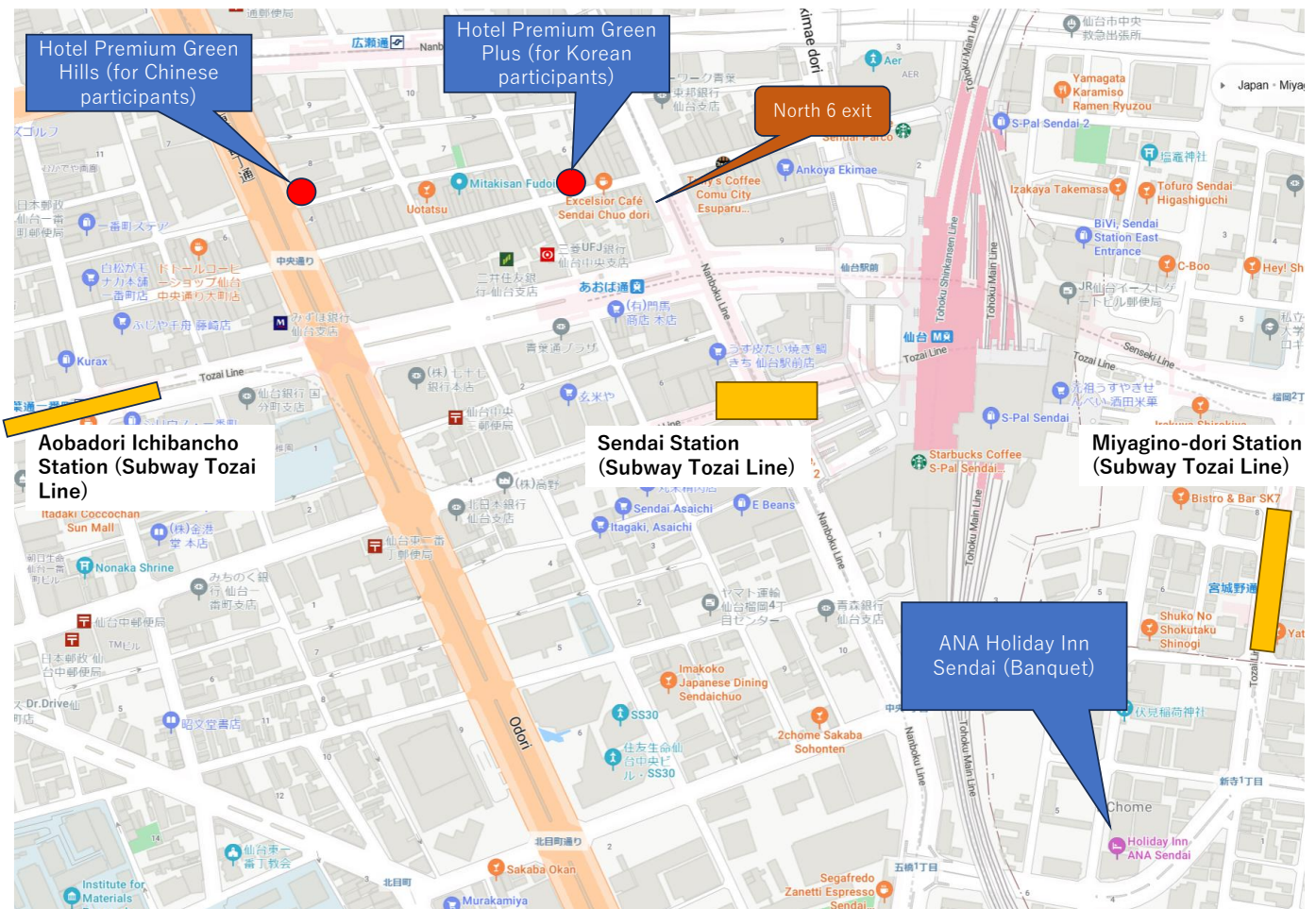
17:05 ~ 17:20

Reflective lens design using perforated dielectrics

Riku KUDO, Hiroyasu SATO, and Qiang CHEN (Tohoku University)

A high-gain reflective Fresnel lens with phase correction by changing the effective dielectric constant using a perforated dielectric has been proposed. Fresnel lens is thin and lightweight and are expected to be applied to 5G communications and millimeter wave imaging. In this report, a perforated dielectric reflective lens with a tilted beam direction by decentering the Fresnel zone and an overlapping lens structure to achieve multi-beam capability were developed and was evaluated by experiments.

Map of Sendai City



Map of Aobayama Campus

