NTU GICE **Newsletter**
Graduate Institute of Communication Engineering, National Taiwan University

Vol. 11, No.1 FEB. 2020 http://www.comm.ntu.edu.tw gicenewsletter@ntu.edu.tw

In this issue

GICE Honors 1

Message from the Director 2

Technology Developed in GICE
- Unsupervised Learning 1-3
- 38-GHz Tx/Rx Beamformers for 5G mm-Wave Phased-Array 4-6

Activities 6-7
- The 2019 Second Semiannual Report of Taiwan Electromagnetic Industry-Academia Consortium on "The Key Technology and Application of 5G Mobile Wideband"

Corner of students News 8

Technology Developed in GICE

Unsupervised Learning

from Data Science and Smart Networking Group

Limitation of Supervised Learning

Thanks to deep learning, the power of AI has grown rapidly. However, the learning capability of machines still falls far behind that of humans. The learning is usually supervised. Typical supervised learning can be simplified as Fig. 1. To learn an AI system, whose input is x (audio or image in the figure), and the output is y (text transcription or image class), we need to provide a large amount of *paired* x and y, (x_1,y_1), (x_2,y_2) … (x_N,y_N), where N is a very large number. For example, to train a speech recognition system today, developers must not only collect a large amount of audio, but also must provide the manual transcriptions corresponding to this audio. Without paired text, there is no way for the machine to learn to recognize speech based on today's technology. For another example, to train an image recognition system, the enormous numbers of images and the corresponding object labels are usually inevitable.

However, to collect this kind of paired data, human annotation is usually needed, which is very expensive. For speech recognition, only rich-resourced languages – for example, Chinese and English – with tens of thousands of hours of paired data can achieve relatively low recognition error rates. For low-resourced languages – for example, Taiwanese Hakka – high-quality speech recognition systems are not available. However, there are more than seven thousand languages in the world, of which more than 95%

(Continued on page 2)

GICE Honors

Prof. Hung-yi Lee

2019 MoST
Ta-You Wu Memorial Award

Prof. Hsuang-Jung Su

2019 MoST Future Tech Demo and Breakthrough Award

Message from the Director

Hsuan-Jung Su

Professor & GICE Director

Happy New Year! We are pleased to share the great news of Prof. Hung-Yi Lee, who was recognized by the MOST with the Ta-You Wu Memorial Award which is awarded to top-notch you researchers. We are also happy to see the teams led by Prof. Hsuan-Jung Su and Prof. Shau-Gang Mao both winning the 2019 MOST Futuristic Breakthrough Technology Award. Congratulations!

In this issue, we invite Prof. Hung-Yi Lee to share his research results on unsupervised learning and its application examples. Prof. Huei Wang also shares his research results on 38-GHz Tx/Rx Beamformers for 5G mm-Wave Phased-Array, which is very important for 5G wireless communications. In the Corner of Student News, Ta-Jung Liu, a student of NTU GICE, shares his internship experience at Facebook, New York. Please enjoy reading this issue and I wish you wonderful Chinese new year holidays.

Technology

(Continued from page 1)

are low-resourced. Therefore, there is not enough paired data to train speech recognition systems for all languages. This complicates the development of speech technology for a new language.

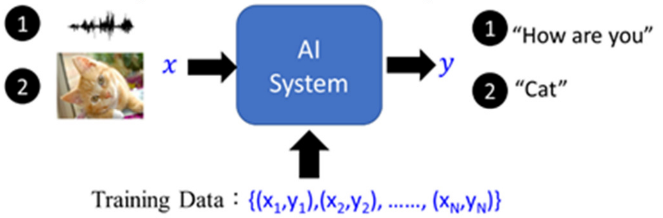


Fig. 1. Typical Supervised Learning.

Unsupervised Learning and Successful Examples

Is it possible for a machine to learn with little supervision, or even without any paired data? That is, without providing paired data, can the machine discover the relation between x and y by itself? Take speech recognition as an example, as illustrated in Fig 2. Can a machine learn to transcribe speech into text from a large amount of *unpaired* audio (x_1, x_2, …, x_N) and text (y_1, y_2, …, y_M)? With unsupervised speech recognition, the machine can automatically learn human language by watching a large number of videos on the Internet, or by listening to human conversation in real life without any formal teaching, in contrast to today's limited training framework.

There are some successful examples of the unsupervised learning scenario developed by Prof. Hung-yi Lee's research group. The research below is mainly based on Generative Adversarial Network (GAN).

- Voice conversion (VC) aims to convert the speech signals from a certain acoustic domain to another while keeping the linguistic content unchanged. Transforming the voice of speaker A (x in Fig. 1 and 2) into speaker B (y) can be considered as a typical example of VC. To achieve that, speakers A and B have to read hundreds of sentences with the same content to teach machine how to transform their voices, which is not practical. Some new VC approaches are proposed [1,2]. Only the audio content of speakers A and B are needed. They do not have to read the same sentences, and they don't even have to speak the same languages.
- The conventional chatbot is in general emotionless, and this is a major limitation of chatbots today because emotion plays a critical role in human social interactions, especially in chatting. Therefore, we wish to train the chatbot to generate responses with scalable sentiment by setting the mode for chatting. This can be achieved by GAN, which transforms the style of chatbot response, for example, from negative (x) to positive (y), without paired data [3]. The techniques mentioned here may be extended to conversational style adjustment, so the machine may imitate the conversational style of someone the user is familiar with, to make the chatbot more friendly or more personal [4].
- Summarization is to generate a summary that describes the core ideas of the document. Based on the typical deep learning based model, by reading documents (x) and their corresponding

(Continued on page 3)

Technology (Continued from page 2)

human-written summaries (y), a machine can learn to generate abstractive summaries. However, to train a summarizer with reasonable performance, in general, millions of paired training examples are needed, which limits the application of the technology. With GAN, a generator learns to shorten an input document, while the discriminator checks whether the generator outputs are readable. In this way, the generator learns to generate summaries with less supervision [5].

- Machine is also possible to learn to recognize speech without paired text. After listening to a large amount of audio, a machine can automatically identify word boundaries and identify those audio segments that correspond to the same word. Eventually, based on the context of these audio segments, the machine can recognize the text of the audio. The developed unsupervised learning approach is very different from that of traditional supervised learning; it may be more closely approximate the way infants learn languages. The world's first unsupervised speech recognition system has been developed [6]. Without paired text, it has achieved a phoneme recognition accuracy of 36% [7].

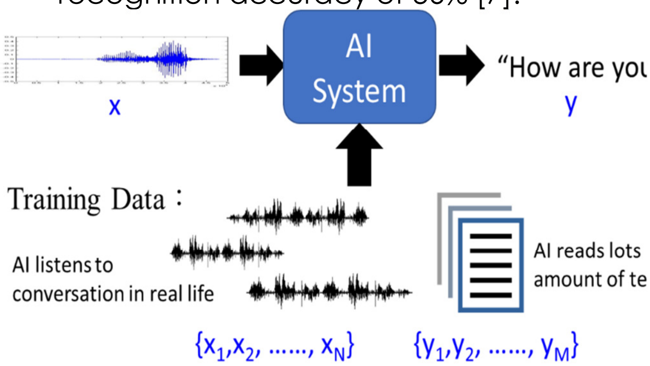


Fig. 2. Unsupervised Learning.

Today's success in deep learning depends on the availability of a large amount of paired data. However, only large companies have the resources to obtain such training data. Due to their small scale, academic institutions or small companies in Taiwan cannot compete with large-scale industry. However, due to the rapid growth of the Internet, unannotated multimedia content is easier to collect now. With the technology developed, the machine will learn with such unpaired data, so that even small and medium-sized enterprises can leverage data from the Internet.

Reference:

[1] Ju-chieh Chou, Cheng-chieh Yeh, Hung-yi Lee, "One-shot Voice Conversion by Separating Speaker and Content Representations with Instance Normalization", INTERSPEECH, 2019

[2] Ju-chieh Chou, Cheng-chieh Yeh, Hung-yi Lee, Lin-shan Lee, "Multi-target Voice Conversion without Parallel Data by Adversarially Learning Disentangled Audio Representations", INTERSPEECH, 2018

[3] Chih-Wei Lee, Yau-Shian Wang, Tsung-Yuan Hsu, Kuan-Yu Chen, Hung-Yi Lee, Lin-Shan Lee, "Scalable Sentiment for Sequence-to-sequence Chatbot Response with Performance Analysis", ICASSP, 2018

[4] Feng-Guang Su, Aliyah Hsu, Yi-Lin Tuan and Hung-yi Lee, "Personalized Dialogue Response Generation Learned from Monologues", INTERSPEECH, 2019

[5] Yau-Shian Wang, Hung-Yi Lee, "Learning to Encode Text as Human-Readable Summaries using Generative Adversarial Networks", EMNLP, 2018

[6] Da-Rong Liu, Kuan-Yu Chen, Hung-Yi Lee, Lin-shan Lee, "Completely Unsupervised Phoneme Recognition by Adversarially Learning Mapping Relationships from Audio Embeddings", INTERSPEECH, 2018

[7] Kuan-yu Chen, Che-ping Tsai, Da-Rong Liu, Hung-yi Lee and Lin-shan Lee, "Completely Unsupervised Phoneme Recognition By A Generative Adversarial Network Harmonized With Iteratively Refined Hidden Markov Models", INTERSPEECH, 2019

For more information please contact:
Professor Hung-Yi Lee
Email: hungyilee@ntu.edu.tw

Technology

38-GHz Tx/Rx Beamformers for 5G mm-Wave Phased-Array

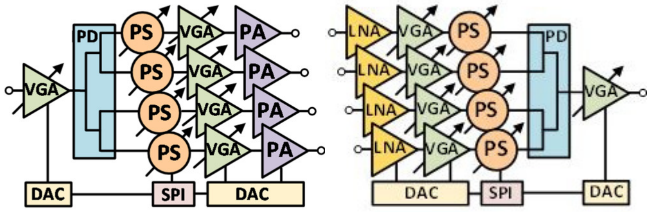
from Electromagnetics Group

INTRODUCTION

The 5G mm-wave applications adopt phased-array systems for EIRP/SNR improvement to overcome path loss issue, and increase the bandwidth for Gbps throughput. Federal Communications Commission (FCC) had released mm-wave bands for 5G including 28/37/39-GHz licensed bands. Many phased-array applications in 28-GHz band had been introduced since this band is going to be the first commercial mm-wave band for mobile communication [1], [2]. On the other hand, 37/39-GHz bands are also attractive for wider available bandwidth [3]. In this manuscript, the 4-element Tx/Rx all-RF beamformer implemented for 5G mm-Wave 37/39-GHz bands phased-arrays are introduced [4].

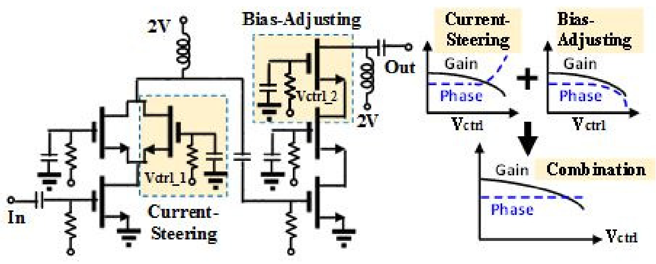
CIRCUIT DESIGN

Fig. 1. shows block diagrams of the Tx/Rx beamformers. The circuits are 4-channels arrays. In Fig. 1(a), the input signal is pre-amplified by the variable gain amplifier (VGA) and split by a 4-way power divider (PD), then the four in-phase signals are fed into Tx chains consisting of phase shifters (PSs), VGAs, and PAs. On the other hand, in Fig. 1(b), the input signals with particular phases are fed into the Rx chains composed of LNAs, VGAs, and PSs. Then four in-phase signals are combined by a 4-way PD and amplified by the VGA. The PS is the combination of a 4-bit (22.5°) switch-type phase shifter (STPS) and a continuously tunable reflection-type phase shifter (RTPS) to improve phase accuracy [5]. The VGA as shown in Fig. 2(a) utilizes current-steering and bias-adjusting technologies [6] to avoid dramatic phase variation in gain tuning. Fig. 2(b) shows the concept of phase compensation. The PA adopts the neutralization technique for stability improvement, and class-AB/class-C amplifiers for linearity and efficiency enhancement (Fig. 3) [7]. The LNA is a three-stage cascade common source amplifier with source degeneration inductors for stability and linearity improvement. All of controlled voltages for gate biases, gain and phase tuning are provided by embedded 6-bit DACs and SPI buses for precise and dynamic voltage control.



(a) (b)

Fig. 1. Block diagrams the proposed (a)Tx, (b)Rx beamformers.



(a) (b)

Fig. 2. (a) Schematic of the VGA. (b) Phase compensation.(a)

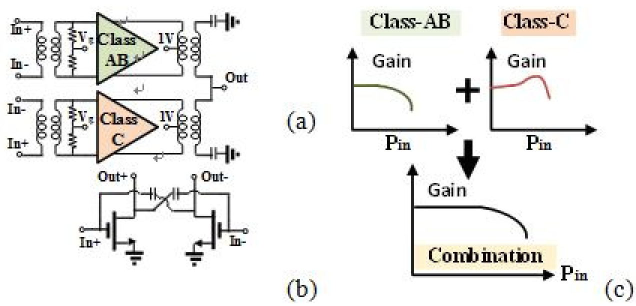


Fig. 3. (a) PA output stage. (b) Neutralization technique. (c) Linearity enhancement.

EXPERIMENT RESULTS

The beamformer are fabricated using TSMC 65-nm CMOS process. Fig. 4. shows chip photos. Chips are on-wafer measured. Fig. 5 shows S-parameter simulations and measurements, the gains of Tx/Rx beamformers are 37-40 dB per channel in 36-40 GHz. Fig. 6(a)-(b) are phases measurements, the PSs achieve 360° phase tuning range with 22.5° resolution. Fig. 6(c)-(d) show 20 dB gain tuning by VGAs. Fig. 6(e) shows root mean square (RMS) gain/phase errors in phase tuning. In 36-40 GHz, RMS gain/phase errors are lower than 2 dB/5.5° (without VGA tuning). The phase compensation by the RTPS appropriately

(Continued on page 5)

Technology *(Continued from page 4)*

calibrates the phase inaccuracy of the STPS. Fig. 6(f) shows ±5° phase variation in 20 dB gain control (without RTPS tuning). The Tx OP_{1dB} at 38GHz is 10-12.3 dBm at the gain of 27-40 dB. The Rx IP_{1dB} at 38GHz is -41 dBm at 40 dB peak gain. The NF per channel of the Rx is 6-9 dB in 36-40 GHz. The measured *IIP3*/*OIP3* of Tx is -18 dBm/20 dBm at 38GHz with peak gain of 38 dB. The measured *IIP3* of the Rx is -33 dBm at 38GHz with 40 dB peak gain. The power consumptions of the Tx/Rx beamformers are 995 mW/693 mW. Moreover, the Tx/Rx beamformers are utilized to implement the 4-element phased-array modules with GaAs PA/LNA and patch antennas for data link as shown in Fig. 7. In Fig. 8, the 38-GHz data-link demonstrates 0.85 Gb/s (64 QAM OFDM, 250M buad rate) with EVM of 5.8% in 4m link distance.

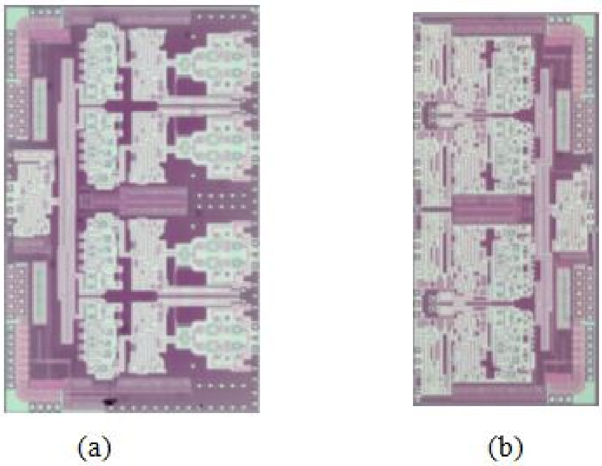


Fig. 4. Chip photographs of (a) Tx (2.3×3.7 mm²) and (b) Rx (1.7×3.6 mm²) beamformers

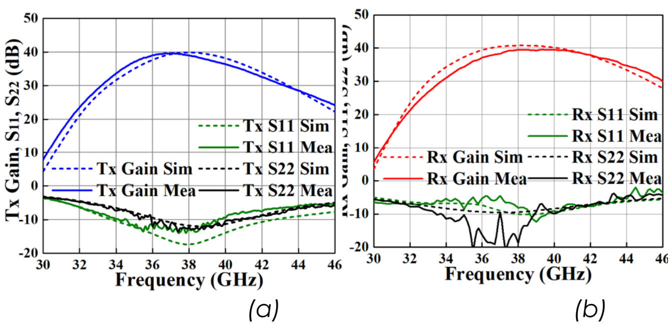


Fig. 5. S-parameter simulation and measurement of (a) Tx and (b) Rx beamformers per channel.

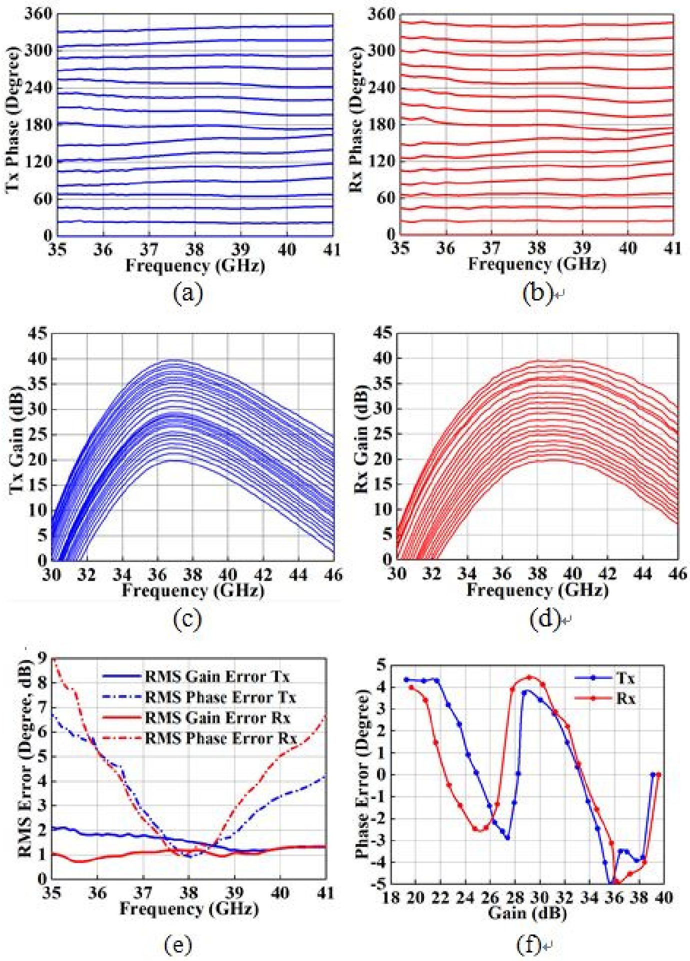


Fig. 6. Tx/Rx beamformer measurements of (a) (b) phase tuning, (c) (d) gain tuning, (e) RMS phase error and gain errors in phase tuning, (f) phase variation in gain tuning at 38GHz.

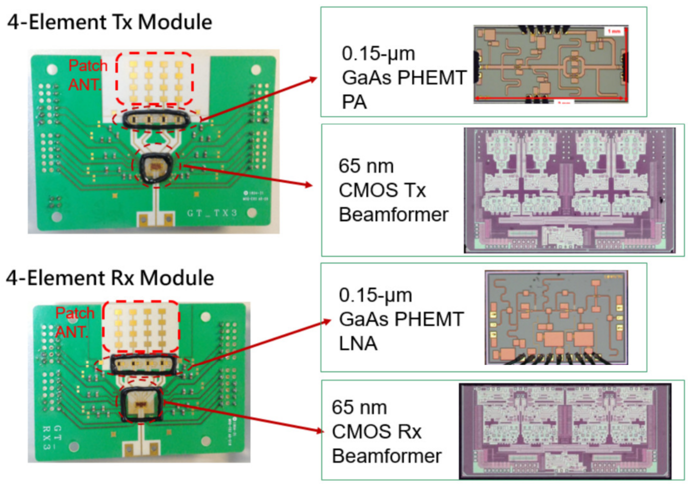


Fig. 7. 4-element phased-array Tx/Rx modules.

(Continued on page 6)

Technology (Continued from page 5)

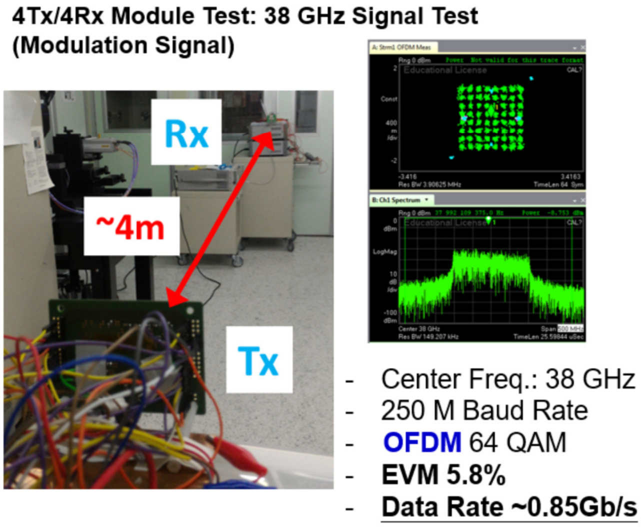


Fig. 8. Wireless data-link test of the Tx/Rx modules.

Conclusion

This manuscript introduced the 38-GHz 4-element Tx/Rx beamformer. The all-RF beamformers showed good performances and the 4-element phased-array modules demonstrate 0.85 Gbps wireless data-links which are suitable for 5G mm-Wave communications.

References

[1] K. Kibaroglu et al, "A Low-Cost Scalable 32-Element 28-GHz Phased Array Transceiver for 5G Communication Links Based on a 2 × 2 Beamformer Flip-Chip Unit Cell," IEEE J. Solid-State Circuits, vol. 53,no. 5, pp. 1260–1274, May. 2018.

[2] H.-T. Kim et al, "A 28-GHz CMOS Direct Conversion Transceiver With Packaged 2×4 Antenna Array for 5G Cellular System," IEEE J. Solid-State Circuits, vol. 53,no. 5, pp. 1245–1259, May. 2018.

[3] Y. Yin et al., "A 37-42 GHz 8x8 phased-array for 5G communication systems with 48-50 dBm EIRP," Proc. IMS, 2019, pp. 480–483.

[4] C.-N. Chen et al., "36-40 GHz Tx/Rx Beamformers for 5G mm-Wave Phased-Array," Proc. APMC, 2018, pp. 756–758.

[5] C.-H. Wu et al, "Design of a K-band Low Insertion Loss Variation Phase Shifter Using 0.18-μm CMOS Process," in Proc. APMC, Nov. 2010.

[6] D.-S. Siao et al, "A 60 GHz low phase variation variable gain amplifier in 65-nm CMOS," IEEE Micro. Wireless Compon. Lett., v01. 24, no. 7, pp. 457-459, Jul. 2014.

[7] H. Wang et al, "A 38-GHz High-Efficiency and Low-Quiescent-Power Power Amplifier for Phased Array Applications in 65-nm CMOS Process," in IEEE MTT-S International Microwave and RF Conference (IMaRC), 2017.

For more information please contact:
Professor Huei Wang
Email: hueiwang@ntu.edu.tw

Activity

The 2019 Second Semiannual Report of Taiwan Electromagnetic Industry-Academia Consortium on "The Key Technology and Application of 5G Mobile Wideband"

The 2019 Second Semiannual Report of Electromagnetic Industry-Academia Consortium was held at Barry Lam Hall, National Taiwan University (NTU) in Taipei, Taiwan, R.O.C., on Friday, Oct. 4th, 2019. The semiannual report was organized by Taiwan Electromagnetic Industry-Academia Consortium, 5G Antenna and RF Technology Consortium Center of Ministry of Education, High-Speed RF and mmWave Technology Center of NTU, Industry Liaison Office of NTU. The co-organizers are The Dept. of Electrical Engineering of Yuan-Ze University (YZU), The Dept. of Electrical Engineering of NTU, Graduate Institute of Communication Engineering of NTU, and 5G RF Industry-Academia Technology Consortium.

The semiannual report was composed of five distinguished sessions. Each topic of the five

Activity

sessions is as follows: The Application Opportunities of 5G Technologies and Unmanned Vehicle, The 5G Technology Development of IoV and Autonomous Vehicle, The Smart IoT Application and Industry Trend of 5G, The Opportunity and Challenges of 5G mmWave Technology, and The Communication Index Design and Measurement of Massive MIMO Antennas.

The outstanding speakers from the industrial and academic fields, including the convener of by Taiwan Electromagnetic Industry-Academia Consortium Prof. Ruey-Beei Wu, Prof. Shau-Gang Mao from NTU, Prof. Jeng-Rern Yang from YZU, Vice Manager Huei-Ru Tseng and Senior Engineer Jun-Xiang Ji from Information and Communication Research Laboratories of ITRI, and General Manager Tsung-Wen Chiu gave the wonderful speeches to the hundreds of listeners.

Between each session, the audiences were having their great time enjoying the delicious food and participating the recruit campaign held by National Chung-Shan Institute of Science and Technology, Microelectronics Technology Inc., Auden Technology Inc. and Unimicron Technology Corporation.

The semiannual report is aim at how to improve the convenience of our life with 5G. We need to solve the problem from the users' standpoint, the social requirements the form of next-generation societies. For example, the unmanned vehicles using the low-latency characteristic of 5G can deliver things to people. Using Internet of Vehicles with 5G to improve road safety and make the traffic flow more efficient than before. Learning about how autonomous vehicles can make our life so much easier. How to improve the coverage of 5G millimeter wave band using small-cell base station. What conditions do we need to consider with 5G when developing millimeter wave technology. How do we design and operate the antennas for millimeter wave application. After making the antennas, how to measure the antennas with accurate results and still be cost-effective.

During the semiannual report, the audiences from the industries and academics were provided with a good chance of broaden their vision on 5G technology. Under an unstoppable development of 5G, including the upcoming commercial operation within next year, the audiences were further realized the position of Taiwan in the 5G global industries.

Corner of student news

Facebook Internship Experience

Article by TA-JUNG LIU

Last fall, I had the privilege to work on Facebook, one of the best tech companies in the world. My office was in NY, a fast-paced while fascinating city. My title was a research intern and mainly focused on developing a brand new part of the ASR (automatic speech recognition) system. The project was exciting to me because it was a very practical problem and had a huge potential production impact in the future. I had a really tough beginning of this project because it was like a project from zero. No similar codebase or project was similar to the direction that I was working on. I encountered numerous research and engineering challenges. Under these circumstances, I found out that communication was very critical. I always kept the manager and mentor knowing what I was currently working on, what challenge I was facing and even whether the plan was still on track or reasonable. At the end of my sixteen weeks internship, I got a promising result and got a return offer, which I could say a happy ending.

This was my first time to intern at such a big tech company and I was amazed by several aspects.
First, no clocking on or off as needed. I could actually stay at home all day, no need to go to the office. Even meeting could be attended remotely. Second, ALL FOOD WAS FREE. There were three restaurants in the company serving breakfast, lunch, dinner and even afternoon tea. Many kinds of snacks and milk were served all day long. Third, there were all kinds of reimbursement, like the company paid the gym for me. However, all the above benefits actually indicated higher pressure. It was like a company to give all the benefits and made it more convenient for our life only to save our time for working, so we were expected to be more productive for sure.

Photo with Statue of Liberty

During the internship in NY, housing was fully supported by the company. It was so cool to live in the center of NY, one of the most expensive places in the world, for free and for such a long time. NY is a melting pot; there are all kinds of people there. After staying there for a long time, there was a feeling that it would not be strange if you belonged there. There are many famous landmarks in NY, like the Statue of Liberty, Grand Central, Brooklyn Bridge, and Times Square. Viewing them in person is just not real because it is like all these scenes should have only been seen in the movie. Overall speaking, I learned a lot, including academic and cultural aspects, from this internship.

National Taiwan University Graduate Institute of Communication Engineering

No.1, Sec.4, Roosevelt Road, Taipei 10617, Taiwan

Phone
+886-2-3366-3075

Fax
+886-2-2368-3824

E-mail
gicenewsletter@ntu.edu.tw

Visit us at:
http://www.comm.ntu.edu.tw

Editor in Chief
Prof. Borching Su

Editor
Chiao Yun Kang