NTU GICE Newsletter
Graduate Institute of Communication Engineering, National Taiwan University
Vol. 7, No.3 AUG 2016 http://www.comm.ntu.edu.tw gicenewsletter@ntu.edu.tw

In this issue

Technology Developed in GICE

Towards Machine Comprehension of Spoken Content

from Communication and Signal Processing Group

Multimedia and spoken content presents more attractive information than plain text content, but they are more difficult to display on a screen and be selected by a user. As a result, accessing large collections of multimedia and spoken content is much more difficult and time-consuming than the latter for humans. In speech processing & machine learning laboratory, we are developing the machines which can automatically understand spoken content and summarize the key information for humans. Below, we are going to highlight three achievements in 2016.

TOEFL Listening Comprehension Test by Machine

We develop and propose a new task of machine comprehension of spoken content, TOEFL listening comprehension test by machine. TOEFL is an English examination which tests the knowledge and skills of academic English for English learners whose native languages is not English. In this examination, the subjects would first listen to an audio story around five minutes and then answer several questions according to that story. Each question has four choices where only one is correct. A real example in the TOEFL examination is shown in Fig. 1. The upper part is the manual transcription of a small part of the audio story. The questions and four choices are listed too. The correct choice to the question in Fig. 1 is choice A. The questions in TOEFL are not simple even for a human with relatively good knowledge. This task is believed to be very challenging for the state-

(Continued on page 2)

GICE Honors

Prof. Lin-Shan Lee
「The 31th Academicians of Academia Sinica」

Prof. Wanjiun Liao
「IEEE ComSoc Board of Governors Member」

Prof. Tzong-Lin Wu
「IEEE EMC Society Board of Directors (BoD) Member」

Message from the Director

Tzong-Lin Wu

Professor & GICE Director

Time flies, seasons switch, but the NTU GICE spirit to nurture elites in the electrical engineering field lingers on.

In this issue, we invite Prof. Hung-Yi Lee and Prof. Hsi-Tseng Chou to share their recent research. Besides, we are honored to declare that Prof. Lin-Shan Lee was awarded as the 31th Academicians of Academia Sinica which is a supreme glory in Taiwan scholarly societies.
Due to the continuous dedication in international collaboration and focusing on students' recruitment, it's rejoicing to see GICE international affairs are on the upswing.
Let's read the sharing article of Robin Jeanty who will be the first one awarded degrees from NTUGICE and INPT-ENSEEIHT following the reciprocal agreement of Double Degree Program.

Technology *(Continued from page 1)*

of-the-art spoken language understanding technologies.

We propose a listening comprehension model for the task defined above, the Attention-based Multi-hop Recurrent Neural Network (AMRNN) framework, and show that this model is able to perform reasonably well for the task. In the proposed approach, the audio of the stories is first transcribed into text by ASR, and the proposed model is developed to process the transcriptions for selecting the correct answer out of 4 choices given the question. The attention-mechanism proposed in this paper can be applied on either word or sentence levels. After training the machine by about 700 questions from the previous TOEFL exams, on manual transcriptions, the proposed model achieved 51.56% accuracy, while the very capable memory network got only 39.17% accuracy. Even on ASR transcriptions with WER of 34.32%, the proposed model still yielded 48.33\% accuracy.

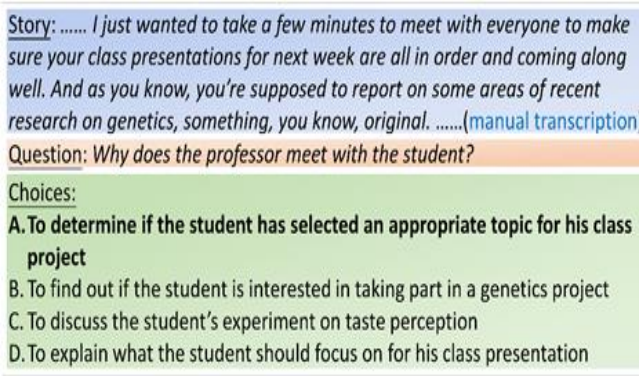


Fig1. An example of TOEFL listening comprehension test. The story is given in audio format, and its manual transcription is shown. The question and choices are provided in text format.

Interactive Spoken Content Retrieval by Deep Reinforcement Learning

User-machine interaction is important for spoken content retrieval. For text content retrieval, the user can easily scan through and select on a list of retrieved item. This is impossible for spoken content retrieval, because the retrieved items are difficult to show on screen. Besides, due to the high degree of uncertainty for speech recognition, the retrieval results can be very noisy. One way to counter such difficulties is through user-machine interaction. The machine can take different actions to interact with the user to obtain better retrieval results before showing to the user. The suitable actions, for example, requesting for extra information from the user, returning a list of topics for user to select, etc., depend on the retrieval status. We use deep reinforcement learning, which is the technology also used in Alpha Go, to determine the machine actions for interactive spoken content retrieval. Deep reinforcement learning bypasses the need for using hand-crafted rules. It is shown to achieve significantly better performance compared with the previous hand-crafted systems. We further found that even with raw relevance scores alone without any human knowledge, we achieved very good performance.

Can machine learn human language from audio story without any supervision?

Today, machine can read a lot of text and

(Continued on page 3)

Technology *(Continued from page 2)*

represent each word (in text) by a vector representations of fixed dimensionality. This technique has been shown to be very useful in many application scenarios, in particular due to the semantic information they carry. We are considering whether machine can offer the vector representations of fixed dimensionality for variable-length audio segments after listening to hours of audio stories. To achieve that, we propose to learn from audio data without human annotation by using Sequence-to-sequence Autoencoder (SA). SA consists of two RNNs equipped with Long Short-Term Memory (LSTM) units: the first RNN (encoder) maps the input audio sequence into a vector representation of fixed dimensionality, and the second RNN (decoder) maps the representation back to the input audio sequence. The two RNNs are jointly trained by minimizing the reconstruction error. Denoising Sequence-to-sequence Autoencoder (DSA) is further proposed offering more robust learning. We found that these vector representations are shown to describe the sequential phonetic structures of the audio segments to a good degree. We have used the vector representations in real world applications such as query-by-example Spoken Term Detection (STD). In this STD application, the proposed approach significantly outperformed the conventional Dynamic Time Warping (DTW) based approaches at significantly lower computation requirements. For the future work, we are considering other application scenarios.

Reference:

[1] Bo-Hsiang Tseng, Sheng-Syun Shen, Hung-Yi Lee, Lin-Shan Lee, "Towards Machine Comprehension of Spoken Content: Initial TOEFL Listening Comprehension Test by Machine", Interspeech 2016

[2] Yen-Chen Wu, Tzu-Hsiang Lin, Yang-De Chen, Hung-Yi Lee, Lin-Shan Lee, "Interactive Spoken Content Retrieval by Deep Reinforcement Learning" , Interspeech 2016

[3] Yu-An Chung, Chao-Chung Wu, Chia-Hao Shen, Hung-Yi Lee, Lin-Shan Lee, "Audio Word2Vec: Unsupervised Learning of Audio Segment Representations using Sequence-to-sequence Autoencoder" , Interspeech 2016

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

Progress of MM-Wave Antenna Technologies for 5G Applications

from Electromagnetics Group

Introduction

Millimeter wave (mmW) [1] technologies are expected to play an important role in 5G to provide broad frequency spectrum and high data rate. In this, the antenna technologies play an important role to overcome the drawbacks of mmW in high propagation loss and sensitive RF characteristic variations. A task force of high gain antenna technologies is formed under the collaboration between National Taiwan University and National Chung-Shan Institute of Science and Technology. The antenna technologies under development include the phased array antennas [2], reflector [3] and reflectarray antennas [4] that are capable of radiating high directive patterns. The application scenarios include dynamic beam steering, beam-switching and multi-beam coverage. The supporting technologies include the calibration of phased array antenna system, measurement technologies, radiation pattern optimization techniques and the characteristics extraction for the integration of RF subsystem. The goal is to establish the required technologies and fulfill the need of industrial applications for the 5G system.

Development Architecture

Fig. 1 illustrates the architecture of development strategy. Multi-port antenna

(Continued on page 4)

Technology (Continued from page 3)

systems are featured to provide the characteristics of diversities for smart antenna and MIMO applications. In the antenna hardware development, two different aspects are examined. The first aspect considers conventional phased array antennas, where RF T/R modules are implemented at antenna ports directly while the second considers the multi-beam strategy by implementing the T/R module at the beam ports to realize the beam-switching or multi-beam cell coverage. In functionality, the first realize the dynamic beam forming for beam steering while the second provides the multi-beam coverage splitting to increase the system capacity. In the second case, the overlapping of beams has to be orthogonal in order to fulfill the need of coverage. The desired gain at the intersection is larger than 20dBi for all cases.

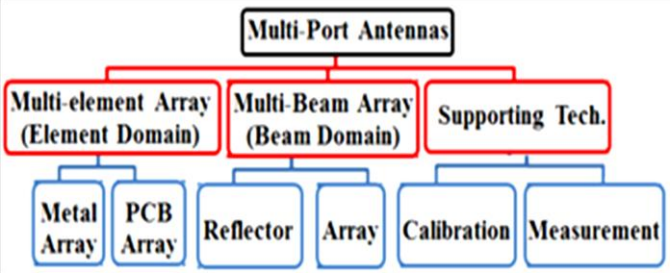


Fig. 1 Implication of MM-wave on 5G scenarios

Due to the RF sensitive characteristics at mmW by its short wavelength, supporting techniques to assure the antenna radiation characteristics are developed including implementing the phase calibration techniques combining the antenna measurement system with the numerical optimization techniques [5]. This technique does not require priori information of phase variation outcome from the RF components including phase shifters. The technique utilizes the optimization techniques to find the control codes of RF components. Other supporting techniques including the grating lobe suppression, which is required in the situation of beam-steering and that uses larger spacing, and RF beam-forming circuit (BFC) self-calibration are also developed.

Current Progress

The frequency is at 38GHz with an estimation of 20dBi requirement for the base station antennas to provide a coverage radius of 50m. In this case, the antenna of user equipment (UE) requires 10dBi.

A. Linear Array of Waveguide-type antennas

In order to fulfill the requirement of radiation efficiency, 16 (8x2) elements of metal waveguide-type antennas are first developed for the horizontal beam steering, where the maximum scan angle of 60 degrees is designed. The waveguide is retained for the operation at fundamental mode with a vertical linear polarization. In this case, the second row of array is shifted by a half period to avoid grating lobes, where the period is 0.65λ. The simulated pattern of array factor for the directional beam pointing toward 40 degrees is shown in Fig. 2. In this case, the projection of elements on the horizontal dimension is less than 0.5λ (only 0.325λ).

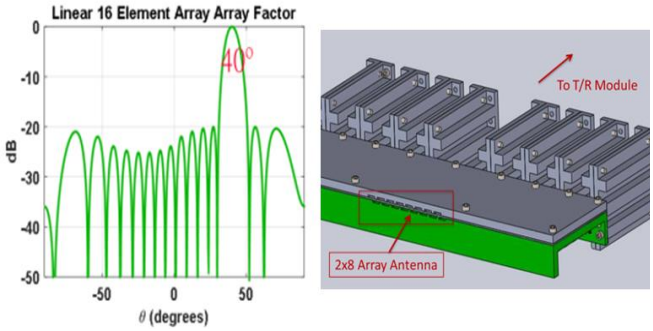


Fig. 2: Radiation patterns of active linear array of waveguide type antennas

B. Illustration of Multi-beam Antenna

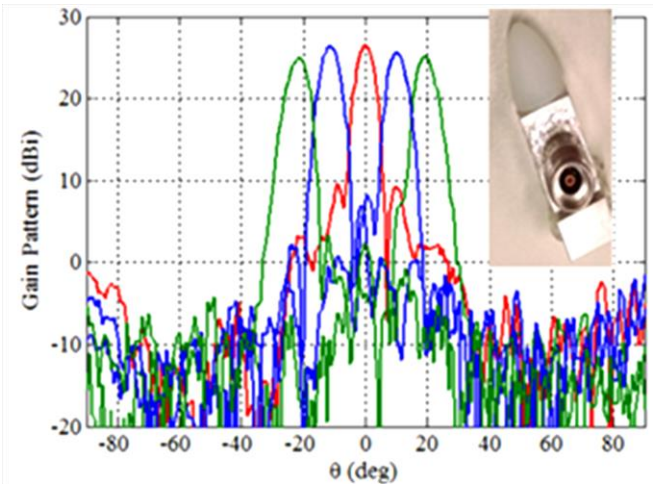
The multi-beam radiation is designed for a sectorized coverage of 50-55 degrees, which provides the advantages of multi-user communication and reducing the frequent beam-steering when the target UE moves or multi-users present at separated locations. Two approaches have been employed by considering metal reflector antennas fed by multiple waveguide type antennas to produce multi-beam radiation, and phased array antennas with a proper BFC. In the reflector design, the horn sections of feeds are replaced by RF lens to reduce the mutual coupling between adjacent feeds. Also the reflector surface is synthesized to tradeoff the gains of different beams [2]. The radiation gain of the focused feed is designed by roughly 26dBi. The location of each feed is aligned to assure that the beam peaks will

(Continued on page 5)

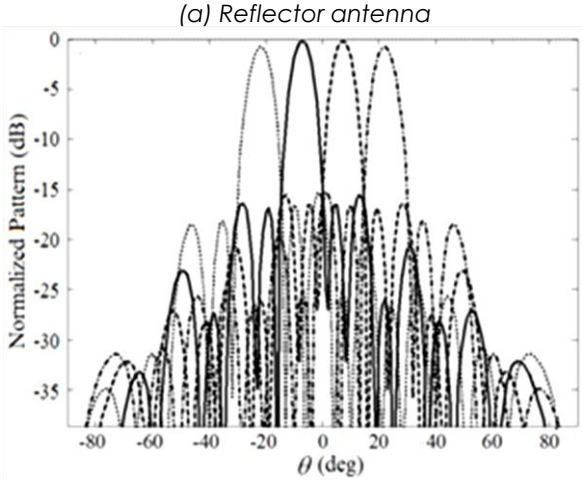
Technology *(Continued from page 4)*

fall into the first null of the adjacent beams as shown in Fig. 3(a), where the reflector has a size of roughly 10cm with gains of roughly 25.5dBi and sidelobe levels of roughly -25dB. Fig. 3(a) shows intersecting point is at nearly 20dBi.

In the second approach of using phased array antennas, the design target is to make the beams orthogonally distributed in a relatively arbitrary coverage area. A promising BFC architecture has been developed to feed phased array of antennas with unequal numbers of antenna and beam ports. Fig. 3(b) shows the radiation patterns of four beams, where four beam ports with 12 antenna elements are considered.



(a) Reflector antenna



(b) Phased array antennas

Fig. 3: Multi-beam radiation patterns of reflector antennas and phased array of antennas.

C. Beam Pattern Calibration of Phase Array Antenna

Due to the short wavelength at mmW, the RF components experience severe phase distortions, which even occur to the phase shifters at the same model, and cause significant pattern distortions. This type of distortion is relatively random as the phase distortion is relatively independent to others. We have established a direct calibration procedure that integrates the optimization technique [3,4] with the antenna measurement system. The procedure adjusts the radiation patterns by using the actual phases and amplitudes output from the RF components in the antenna systems. The architecture of the calibration strategy is illustrated in Fig. 4 [6].

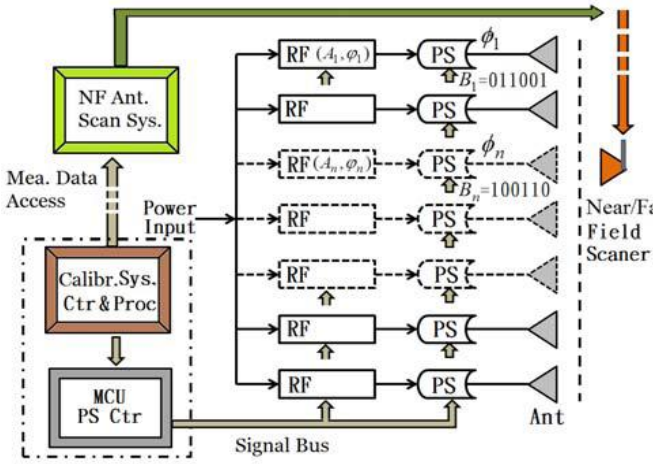


Fig. 4 The architecture for the integrated phase cal. System

The radiation beams are formed based on the applications of GA [7], which is implemented in the processing unit. The variables in optimization are the digital control codes that are binary to control the changes of phase shifters and RF devices (especially the attenuators). Thus the phase shifters and RF devices behave like mathematic transformers to transform the binary codes into phases and amplitudes to excite the array.

Conclusion

The antenna technologies have potentials to enhance the mmW communication performance. Significant progress has been achieved. Subsequent development of integration into RF subsystem is undergoing to form a 5G platform, and will be reported in the near future.

(Continued on page 6)

Technology (Continued from page 5)

References

[1] Rappaport, Theodore S., et al. "Millimeter wave mobile communications for 5G cellular: It will work!." Access, IEEE 1 (2013): 335-349.

[2] L. Schulwitz and A. Mortazawi, "A compact dual-polarized multibeam phased-array architecture for millimeter-wave radar", IEEE Trans. Microw. Theory Tech., vol. 53, no. 11, pp. 3588-3594, 2005

[3] Bankov, S. E., et al. "Design and experimental investigation of a multibeam integrated reflector antenna of the millimeter wave band." Journal of Communications Technology and Electronics 59.6 (2014): 504-522.

[4] D. M. Pozar, S. D. Targonski, and H. D. Syrigos, "Design of millimeter wave microstrip reflectarray", IEEE Trans. Antennas Propag., vol. 45, no. 2, pp. 287-295, 1997

[5] Hsi-Tseng Chou, Sheng-Ju Chou, Chih-Wei Chiu and Tsai-Wen Hsiao, "Shaped Reflector Antenna Design for Orthogonal Multibeam Radiations at Millimeter Waves, 2016 Asia-Pacific International Symposium on Electromagnetic Compatibility & Signal Integrity and Technical Exhibition

[6] H.-T. Chou, D.-Y. Cheng and N.-W. Chen, "The Pattern Calibration of Phased Array Antennas via the Implementation of Genetic Algorithm with Measurement System", IEEE International Symposium on Antennas and Propagation, June, 2016

[7] Y. Rahmat-Samii and E. Michielssen, Electromagnetic Optimization by Genetic Algorithms, 1999 :Wiley

For more information please contact:
Professor: Hsi-Tseng Chou
Email: chouht@ntu.edu.tw

Activities

- GICE Delegates Visited ENSEEIHT and Paris-Sud for Students Recruitment

The École nationale supérieure d'électronique, d'électrotechnique, d'informatique, d'hydraulique et des télécommunications (ENSEEIHT) is a prestigious French engineering school in Toulouse, and established a double degree program with NTU GICE (Graduate Institute of Communication Engineering) in October, 2014. Soon in the summer of 2015, an M1 student of ENSEEIHT, Mr. Robin Jeanty, joined this double degree program under the advisory of Prof. Shih-Yuan Chen. To further strengthen the international collaboration and recruit more double degree students from ENSEEIHT, NTU delegates once again visited ENSEEIHT in January 2016. The NTU delegates include Professor Tzong-Lin Wu, Director of NTU-GICE, Professor Kwang-Cheng Chen, Former Associate Dean of EECS College, and Professor Shih-Yuan Chen.

On January 11th 2016, NTU delegates visited ENSEEIHT to meet the professors and students. In the morning session, Prof. Kwang-Cheng Chen introduced the research activities in NTU GICE Communication and Signal Processing Group while Prof. Shih-Yuan Chen introduced the research activities in NTU GICE Electromagnetics Group. After the luncheon, an introductory presentation is given by GICE Director Tzong-Lin Wu to provide overview of NTU, NTU-GICE, the double degree program, and a new exchange program, which will soon be signed by Presidents of both universities. Surprisingly, more than 30 students attended this event. Some of them have shown great interest in joining the double degree program, conducting research internship in NTU-GICE, or even pursuing Ph.D. degree in NTU-GICE.

(Continued on page 7)

Activities

The headquarter and assembly lines of the European aircraft company Airbus are situated in the north part of Toulouse. On January 12th 2016, Prof. D. Andreu, the International Affairs Head for ENSEEIHT, arranged a private tour in Airbus for the NTU delegates. The NTU delegates first had a guided tour on the assembly chain of the latest model of jet airliner, A380, and then visited the Aeronautical Museum.

Figure 1. Group photo after NTU-ENSEEIHT meeting between professors and students.

Figure 2. Photo of the introductory presentation given by GICE Director Tzong-Lin Wu.

On January 13th 2016, NTU delegates visited University of Paris-Sud. University of Paris-Sud is merged from University of Paris XI and some other premier French engineering schools. In the near future, more prestigious universities and engineering schools will be further merged into a mega school, Universite Paris-Saclay. Prof. Said Zouhdi, Head of International Relations hosted our visit to University of Paris-Sud. First, we had a meeting with professors there to discuss the details of the double degree program between NTU-GICE and University of Paris-Sud, which will soon be signed by both sides. Then, an introductory presentation is given by GICE Director Tzong-Lin Wu to provide overview of NTU, NTU-GICE, and the double degree program. About 16 students attended this event. Some of them have shown great interest in visiting Taiwan for research internship and also joining the double degree program. This visit has been a good start, and we will keep this momentum to establish a long-lasting and fruitful collaboration between the two universities.

Figure 3. Group photo in Aeronautical Museum.

Figure 4. Photo of the introductory presentation given by GICE Director Tzong-Lin Wu.

Figure 5. Group photo after NTU-Paris-Sud meeting between professors and students.

Corner of Student News

Article by Robin Jeanty

Robin Jeanty comes from France and he is pursuing Master Degree by joining double degree program via a collaboration agreement signed by NTUGICE and INPT-ENSEEIHT.

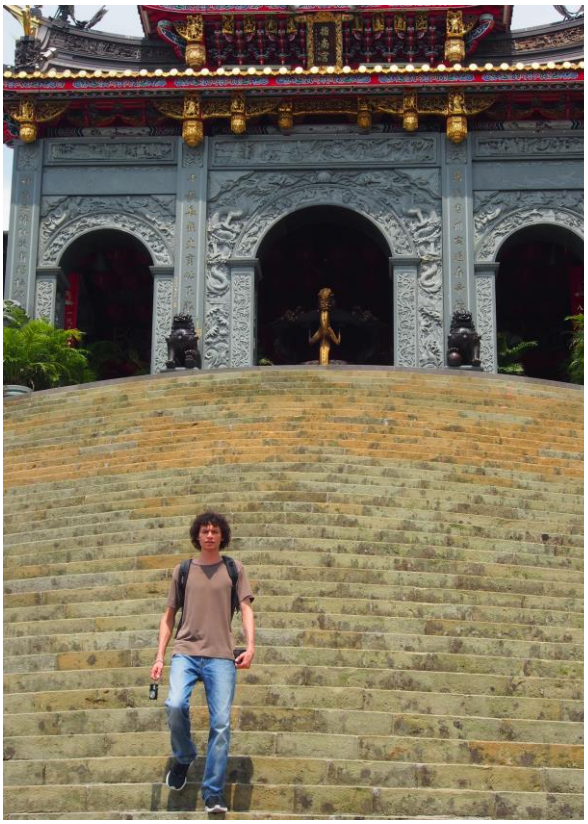
I remember when two years ago, while I was a second-year student in my engineering school in Toulouse (INP-ENSEEIHT), I had to choose where to go for my final year of studies. I had many classical possibilities such as North and South America or Europe which are often chosen by students. However, a few exotic destinations were also available. One of them was Taiwan.

It was a brand new agreement between INP-ENSEEIHT and NTU-GICE signed earlier that year which hadn't had been experienced by any student yet. Therefore, there was no real feedback to be found. I only knew a few facts about Taiwan, namely "French students who went there for an internship came back really delighted", "NTU is a top university in Electrical Engineering"... Nothing truly useful when picturing yourself in a foreign country. Moreover, my mastery of Chinese was completely nonexistent. There was so much unknown that I had to give it a try.

One year ago, I landed in Taipei. The first thing that hit me was Taiwan's weather. Trust me, Wikipedia's "hot and humid climate" is definitely a euphemism. The proper word to describe this climate is sauna and the one and only remedy to survive the heat is probably to zigzag from one convenient store to another and enjoy their air conditioning. Among all the "cultural differences" that Taiwan features, the climate is the only one that I am still struggling with.

For my Master degree, I decided to join a Lab focusing on antennas where I had the chance to meet with a group of very lively and hardworking persons who have always been of great support. At dinnertime, local friends are a real asset to find unexpected places to eat genuine Taiwanese food or just to help you analyze menus written in Chinese. Actually, I have taken Chinese courses since my arrival so I am now able to use some survival Mandarin. However, recognizing the character for "pork" on the menu does not give you the exact part of the animal that you are about to eat. Sometimes, I had to deal with some surprising meals but always tasty.

Now, I am more familiar with Taiwan I can assure that it is a great place to live mostly thanks to its lovely people. Astonishing activities, temples or landscapes (sea shores, mountains) are easy to reach through the different means of transportation. Studying in Taiwan is also a chance to meet people from all over the world and share each other's culture which, in my views, is always a true personal fulfillment. For those reasons, I might extend my stay here. Well, I will for sure.



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. Hung-Yu Wei

Editor
Chiao Yun Kang