Symposium of Yotta Informatics Research Platform for Yotta-Scale Data Science 2019

Abstracts

Program

Opening Remark 10:30 – 10:40

Invited Speakers

10:40 – 11:20

Ryusuke Hayashi (National Institute of Advanced Industrial Science and Technology)

Exploring visual system using deep neural network

11:20 – 12:00

Daniel E. Acuña (Syracuse University)

The effect of innovation on future impact of scientific grants

Lunch Time 12:00 – 13:30

Progress Reports Part 1

13:30 – 13:45

Takaki Sato and Yasumasa Matsuda

Spatial analysis of the degrees of satisfaction and health in Japan

13:45 – 14:00

Taira Nakajima

Supporting system for teaching improvement through reflection using image analysis data of lecture recording: annual report

14:00 – 14:15

Shinichiro Omachi

Historical Document Retrieval by Image Generation

14:15 – 14:30

Akinori Ito

Spoken term detection from speech database of zero-resource language based on DNN-based speaker-independent features

Invited Speaker

14:30 – 15:10

Kazuhisa Takemura (Waseda University)

Escaping from bad decision: How to integrate values of information

Break 15:10 – 15:30

Progress Reports Part 2

15:30 – 15:45

Jiro Gyoba

Human Cognitive Processing during Pattern Selection (Part 2)

15:45 – 16:00

Hideki Kozima

Granular informatics for modeling cognitive styles in ASD and TD populations

16:00 – 16:15

Satoshi Shioiri

Estimation of human judgments of food images

Invited Speaker

16:15 – 16:55

Laurent Itti (University of Southern California)

Attention and saliency in eye tracking with application to detecting disease

Poster Session 17:00 – 18:30

Reception 18:40 – 20:00

Invited Speakers

10:40 – 12:00

Exploring visual system using deep neural network

Ryusuke Hayashi

Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology (AIST)

We can visually recognize a variety of objects and its materials despite substantial appearance variation due to the change of viewing conditions such as noise, lightning and viewing angle etc. A promising approach to elucidate this complex and inherently non-linear visual mechanism is to construct artificial neural networks that recognize objects and materials as humans do and analyze its visual representation in comparison with those of humans/animals obtained from psychophysical and electrophysiological experiments. In my talk, I will introduce three research topics related with this approach.

First, we will report the similarity in visual representations between the higher-order layers of a convolutional neural network (CNN) trained for object classification and the inferior temporal (IT) cortex of monkey brain, a brain area crucial for visual object recognition (Hayashi and Nishimoto, SfN 2013; Hayashi and Kawata, IEEE SMC 2018). In these studies, we recorded neural activity using microelectrode arrays implanted on the surface of the IT cortex. We then tested how accurately the visual features of viewed images in the CNN were predictable from neuronal data and visualized what information was restored by inputting the predicted features into a deep generator network pre-trained to reconstruct photo-realistic images.

Second, I will demonstrate that a neural network consisting of CNN and deep decoder/generator, which was trained by natural images/movies in an unsupervised fashion with information maximization rule, also represents progressively complex visual features as ascending the hierarchy of the CNN in very similar way as human/monkey brain.

Third, I will introduce a collaboration project (Liu et al., VSS 2018) that aimed to understand the visual mechanism of material perception in human. In this project, we developed a CNN that can classify material categories of Flickr Material Database (FMD; Sharan et al., 2013) as accurately as humans, then compared the performance of humans and CNNs in the aspect of noise tolerance. In order to achieve human-level noise tolerance, we need to either apply fine-tuning using the mixed dataset of original and noised images or add a de-noising network to the original material-recognition CNN in a cascaded manner. Although behavioral similarity does not imply computational similarity, these findings provide novel insights into human visual mechanism for material recognition.

The effect of innovation on future impact of scientific grants

Daniel E. Acuña
Syracuse University

Abstract:
Governments and foundations tend to support innovative projects that are too risky for private investment. Evidence supporting the link between such innovation and the eventual impact of a grant is relatively lacking. In this talk, I will present my latest research on measuring innovation in grants and their impact in future productivity and impact. In particular, we analyze more than 130,000 grants funded by the National Science Foundation (NSF) and the National Institutes of Health (NIH) from 2008 and 2015. We use novelty detection techniques from machine learning to quantify innovation and relate that measure to citations of the articles associated with those grants. Across all divisions in NSF and institutes in NIH, and controlling for journal prestige, PI experience, and grant amount, we found that novelty is significantly related to citations. Specifically, we estimate that a fully novel grant would receive 50% more citations than a fully incremental grant. Our results provide strong evidence that the innovation-first policy policies from NSF, NIH, and many other funding agencies well-guided from an impact perspective.

Biography:
Daniel E. Acuña is an Assistant Professor in the School of Information Studies at Syracuse University, Syracuse, NY. The goal of his current research is to understand decision making in science—from helping hiring committees to predict future academic success to removing the potential biases that scientists and funding agencies commit during peer review. To achieve these tasks, Dr. Acuna harnesses vast datasets about scientific activities and applies Machine Learning and A.I. to uncover rules that make publication, collaboration, and funding decisions more successful. Simultaneously, he has created tools to improve literature search (https://scholarfy.net), peer review (http://pr.scienceofscience.org), and detect scientific fraud. He has grants from NSF, DDHS, and DARPA and his work has been featured in Nature Podcast, The Chronicle of Higher Education, NPR, and the Scientist.

Before joining Syracuse University, Acuna studied a Ph.D. in Computer Science at the University of Minnesota, Twin Cities and was a postdoctoral researcher at Northwestern University and the Rehabilitation Institute of Chicago. During his graduate studies, he received a NIH Neuro-physical-computational Sciences (NPCS) Graduate Training Fellowship, NIPS Travel Award, and a CONICYT-World Bank Fellowship.

Progress Reports Part 1

13:30 – 14:30

Invited Speaker

14:30 – 15:10

Escaping from bad decision: How to integrate values of information

Kazuhisa Takemura

Department of Psychology, Waseda University

Bad decisions are made even in serious situations such as selecting a personal career or selecting an important policy in management and politics. In this talk, I will first introduce the conceptual and mathematical frameworks for multi-attribute decision-making and explain from a theoretical point of view, why it is almost impossible to make the best decision. We will assume a set of circumstances related to purchasing a product at a store. Presume, for instance, that we purchase a digital audio player. Consumers make a purchase decision after comparing multiple attributes, such as the prices, number of recordable tracks, sound performance, and designs at stores, or using catalogs. Such decision making after considering multiple attributes is known as multi-attribute decision-making. Multi-attribute decision-making is presumably performed by obtaining various types of information. I would like to exemplify the perspective of form that, when multi-attribute decision-making is viewed from the perspective stated above, it satisfies the rationality standards such as transitivity and connectivity, and conditions considered appropriate in multi-attribute decision-making contradict. This can be derived by the application of and re-interpreting the mathematical structure of the general possibility theorem of group decision-making presented by K. J. Arrow (1951) to the multi-attribute decision-making defined above. Applying Arrow's general possibility theorem on this assumption results in the finding that rational decision-making is possible only when it is based on one-dimensional standards, which suggests that rational decisions generally can not be made if the pluralistic values cannot be ranked, which means that making the best decision would also be meaningless.

I will then give some examples of bad decisions that were determined as bad decisions in experimental studies in both individual and group settings. In experimental studies, people tended to make bad decisions even in fatal situations if they focused on the trivial aspects of a problem. Imagine that you are going to make a choice between two options. One is fresh but cheap steak, and another is very expensive and famous brand steak but expired consumption period. The former is not good but not dangerous, but the latter seems to be very delicious but somewhat dangerous. In this case, choosing the latter option might be bad decision. One hundred and forty two university students were recruited to complete questionnaire of risk judgment. The judgment task was to choose one of two multi-attribute problem (one is fatal and a very dangerous option (Bad decision) and the another is not very dangerous and not a fatal option (not Bad decision)). The bad alternative consists of very risky and fatal information but very attractive information (such as delicious food). For example, Matsuzaka beef is well known as very delicious food in Japan. Fukushima beef is not well known in Japan and people might remember Fukushima nuclear accident in 2011. However, eating raw liver meat is considered more dangerous and fatal than eating heated beef. Surprisingly, many of the participants prefer bad alternatives rather than the better alternatives. More interestingly, bad decisions are not very related to

educational background.

Many social psychological studies have indicated that group decisions are distorted by social pressure such as conformity and authority. However, the implicit assumption of those studies was that each individual is capable of making rational judgments and decisions unless there exists such social pressure. We examined this implicit   assumption. The participants were 64 university students. They were asked to watch three types of the group decision drama in random order. The first condition was the purpose deviation that output of group decision deviate from the original purpose. That is, the purpose was to choose the place for the group tour, but the final decision was to choose good mineral water brand for the members, and was not related to the tour. The second condition is the procedural justice condition that the member tried to keep procedural justice for decision process but the final decision was not supported by the majority. The members firstly decide to construct decision procedure (Voting rule: ballot paper should be folding into two, other folding manner is not valid), and then they made decision but many of the voting were not valid. The final decision was apart from the majority preference. The third condition was the purpose consistent condition in which the final decision was made in line with their purpose. The participants were asked to rate the desirability and adequateness of the discussion procedure and adequateness of discussion procedure. Interestingly, the ratings of procedural justice were not different from those of the purpose consistent condition. The procedural justice condition seems to be just keeping the folding rule, but the final decision was not represented by the majority opinions. Majority of ratings were apart from the better decision. Moreover, more than half of the participants ranked the procedural justice condition as the most desirable.

In addition to providing a psychological model of bad decisions in multi-attribute situations, I offer some suggestions based on empirical research and computer simulation studies on how to avoid making bad decisions. In order to avoid bad decisions, the computer simulation and the experimental findings that I introduced might be useful. According to the findings, the most important thing to avoid bad decisions is to integrate various values of information, and to consider the most important value (which is very related to human value) in the decision situations. If this is ensured, the decision makers can avoid making bad decisions even if they use simple heuristics that had been considered as an important factor of irrational decision by Kahneman and Tversky's classical behavioral economic theory. Considering the most important property seems to be very easy. However, histories all over the world and the empirical studies suggest that this was not very easy. For example, in natural disaster, the most important thing is to save people's lives. This has been forgotten in many situations. In order to not forget the most important value, decision makers should make room for considering the purpose or most important value of the decision problem.

Progress Reports Part 2

15:30 – 16:15

Human Cognitive Processing during Pattern Selection (Part 2)

Jiro Gyoba

Department of Psychology, Graduate School of Arts & Letters, Tohoku University

In my experiment, participants were asked to be an archaeologist who found a set of ancient character patterns consisted of five dots in a virtual ruin. They were permitted to bring only a limited number of patterns, because of the law for the protection of cultural properties. My previous experiment showed that when the limited number is small (6 out of 48), the participants tended to select redundant and regular patterns that are stable for various cognitive transformations. In other words, they selected the patterns of low cognitive processing loads. When the limitation was moderate (12 out of 48), the participants seemed to select also irregular patterns representing various types of cognitive transformations. In my recent experiment, participants rated affective properties of the patterns by Osgood's semantic differential method, after doing the same selection task. As results, the participants chose mainly simple and likable patters when the selection limitation was severe, while they also selected more complex and even unlikable patterns when the limitation was moderate. These findings suggest that humans conduct information triage by adjusting cognitive and affective preferences depending on the degree of limitation.

Estimation of human judgments of food images

Satoshi Shioiri [a, b]
[a] Advanced Institute of Yotta Informatics center, Tohoku University
[b] Research Institute of Electrical Communication, Tohoku University

The information generated has been increasing rapidly and prioritizing data is becoming more and more important. For pictures, preference judgments are factors for prioritization. To help people to make decisions on preferences based on human judgments of image qualities such as tastes, appearance and so on, we attempted to find the relationship between the human judgment scores and features analyzed in a trained convolutional neural networks for object recognition. Based on the relationship, we can develop a computer model that predicts human judgments of image qualities. We used 760 lunch box images with six judgments (deliciousness appearance, preference to eat, preference to arrangement, beautifulness, made for men/women, made for old/young) by 2120 informants (Matsubara and Wada, 2018). We used the pre-trained AlexNet and VGG19 for three types of trainings: (Retraining) parameters totally changed with our data set, (Fine-tuning) training with the data set with trained parameters as initial values, (Frozen layers) the parameters remain unchanged. We found no convergence with VGG19 so we report the results of AlexNet. The estimation for test images showed the best performance for Fine-tuning. Based on small data as 760 images, human judgments can be estimated with accuracy of 69% (38% for the Retaining and 63% for the Frozon layers). We also examined the possible effects of eye movement for seeking better predictions by weighting input images with fixation maps. With smaller number of subjects, we measured gaze points while making each of the same six judgments. Pooling gaze points of all subjects to make a fixation map for each judgment, which is used for weighting areas of images. Although gazes should be on regions of interest for judgments and therefore we predicted that weighting the images by fixation map should focus on features related to judgments, the results did not improve predictions. Further studies are required to investigate the effect of gaze points.

Invited Speaker

16:15 – 16:55

Attention and saliency in eye tracking with application to detecting disease

Laurent Itti
University of Southern California

Eye movements have often been suggested to provide a "window onto the mind". Here, we present new results combining computational modeling with eye tracking to detect neurodevelopmental and neurodegenerative conditions. We first develop a model of visual attention based on the notion of visual saliency, a measure of which stimuli in the visual world may more strongly attract an observer's attention. We then use this model to quantitatively measure an individual brain signature of different individuals. The signature is based on how their patterns of eye movements while they watch 5 minutes of television can be quantified by the saliency model. We apply this technique to several disorders which have traditionally been difficult to diagnose, because there is no reliable biomarker of the disease (Attention Deficit Hyperactivity Disorder, ADHD; Parkinson' disease, PD; and Fetal Alcohol Spectrum Disorder, FASD). The largest group studied was FASD. FASD is one of the most common causes of developmental disabilities and neurobehavioral deficits. Despite the high-prevalence of FASD, the current diagnostic process is challenging and time- and money- consuming, with underreported profiles of the neurocognitive and neurobehavioral impairments because of limited clinical capacity. Participants with FASD and age-matched typically developing controls completed up to six assessments, including saccadic eye movement tasks (prosaccade, antisaccade and memory-guided saccade), free viewing of videos, psychometric tests, and neuroimaging of the corpus callosum. We comparatively investigated new machine learning methods applied to these data, towards the acquisition of a quantitative signature of the neurodevelopmental deficits, and the development of an objective, high-throughput screening tool to identify children/youth with FASD. Our method provides a comprehensive profile of distinct measures in domains including sensorimotor and visuospatial control, visual perception, attention, inhibition, working memory, academic functions, and brain structure. We also showed that a combination of four to six assessments yields the best FASD vs. control classification accuracy; however, this protocol is expensive and time consuming. We conducted a cost/benefit analysis of the six assessments and developed a high-performing, low-cost screening protocol based on a subset of eye movement and psychometric tests that approached the best result under a range of constraints (time, cost, participant age, required administration, and access to neuroimaging facility). Using insights from the theory of value of information, we proposed an optimal annual screening procedure for children at risk of FASD.

Bio:

Laurent Itti received his M.S. degree in Image Processing from the Ecole Nationale Superieure des Telecommunications (Paris, France) in 1994, and his Ph.D. in Computation and Neural Systems from

Caltech (Pasadena, California) in 2000. He has since then been an Assistant, Associate, and now Full Professor of Computer Science, Psychology, and Neuroscience at the University of Southern California. Dr. Itti's research interests are in biologically-inspired computational vision, in particular in the domains of visual attention, scene understanding, control of eye movements, and surprise. This basic research has technological applications to, among others, video compression, target detection, and robotics. Dr. Itti has co-authored over 150 publications in peer-reviewed journals, books and conferences, three patents, and several open-source neuromorphic vision software toolkits.

Poster Session

17:00 – 18:30

Storing perceptually meaningful surround sound information as stereo signals

Jorge Trevino, Shuichi Sakamoto and Yôiti Suzuki
Research Institute of Electrical Communication and Graduate School of Information Sciences, Tohoku University

The sense of hearing is one of the main mechanisms through which we acquire information. Speech carry semantic meaning which may be stored efficiently as text. Other features of sound, such as pitch, are important in order to preserve rich contents like a singer's performance. Beyond these, humans use their two ears to infer the positions from where sounds originate. This ability is known as spatial hearing and a traditional way to store spatial sound information is as multi-channel recordings. Since humans have two ears, two audio channels appear to be enough; however, we normally move and rotate our heads while listening to sounds. This is known as active listening. Storing spatial sound information compatible with active listening requires dozens or hundreds of audio channels. Two examples are the spherical microphone array recordings used in the SENZI algorithm for headphone presentation, and the spatial encodings used in high-order Ambisonics (HOA). Spatial sound can consume extremely large amounts of storage, limiting the number of contents that can be preserved. The present research attempts to lower the multi-channel requirements by imposing a perceptually meaningful assumption: humans cannot separately localize two sounds which fully overlap in the frequency domain. That is, when presented with two sound sources that encompass the same frequencies, we perceive them as a single source. This phenomenon has been used in stereo sound systems to present sounds between the two loudspeakers. In particular, our proposal is to downmix HOA spatial encodings by dividing them in frequency bands and assigning two values to each band: a total energy and a direction of incidence. The two parameters are stored using only two audio channels, encoded as level and phase differences. The result is a stereo-compatible two-channel signal that can be decoded into full surround HOA signals carrying rich spatial sound information.

Brain substrates underlying a perceptual interaction of vision and olfaction.

Yinan Jiang, Shuta Maekawa, Kaiji Yamamichi and Nobuyuki Sakai
Graduate School of Arts and Letters, Tohoku University

In our daily lives, we do not identify odors only by the olfaction but with the vision. We are investigating how this polymodal perception occurs. This study aimed to compare effects of the colors and the shapes of visual stimuli on perception of the olfactory stimuli. This study also aimed to compare the brain responses to odor-visual congruent stimuli and to incongruent stimuli.

Ten university students participated in this study. The participants were asked to evaluate the intensity and the hedonics of the odor presented with a picture. During evaluation, brain responses of the participants were recorded with the near infra-red spectroscopy (FOIRE-3000, Shimadzu). Strawberry flavor and lemon flavor were used as the olfactory stimuli. Pictures of strawberry and lemon were used as the visual stimuli, and the half of the pictures colored in red and the other half in yellow. Thus the number of the stimuli was 8 (2 odors x 2 colors x 2 shapes). The odor-visual stimuli were divided into two groups; the congruent stimuli and the incongruent stimuli. The strawberry flavor with red strawberry picture and the lemon flavor with yellow lemon picture were congruent stimuli. On the other hand, the incongruent stimuli were divided into sub groups. For example, the strawberry flavor with yellow lemon picture as vision-incongruent stimuli, the strawberry flavor with red lemon picture as shape-incongruent stimuli, the strawberry flavor with yellow strawberry stimuli as color-incongruent stimuli, and so on.

Participants evaluated the odor most pleasant and strongest when the odor were presented with the congruent stimuli. The behavioral tendency showed the congruent > the shape-incongruent > the color-incongruent > the vision-incongruent stimuli. This result suggested that the shape is dominant character in the odor-visual interactions.

Attention alters the pattern of recalibration of vocal-auditory subjective synchrony

Kosuke Yamamoto[1], Hideaki Kawabata[2]
1. Research Institute of Electrical Communication, Tohoku University
2. Department of Psychology, Keio University

Prolonged exposure to temporal gaps within multisensory stimuli can induce a recalibration of the subjective synchrony of audiovisual and sensorimotor modalities. While selective attention towards the temporal structure or non-temporal features of stimuli is believed to modulate the recalibration pattern in the audiovisual domain, it is unclear what effect selective attention has on the sensorimotor domain. Thus, we examined how selective attention to temporal and non-temporal information modulates subjective synchrony in vocalization by implementing judgment tasks for stimulus features, synchrony, and temporal order during exposure to vocal-auditory lag. We found that exposure to lag with synchrony-oriented attention shifted the point of subjective synchrony in the opposite direction of the lag; in contrast, exposure to lag with order- and feature-oriented attention induced normal temporal recalibration.

Enhancement of Perceived Reality by Body Vibration Adapted to Foreground Components of Audio-Visual Contents

[1,2] Shota Abe, [2] Zhenglie Cui, [2] Shuichi Sakamoto, [2] Yôiti Suzuki, and [3] Jiro Gyoba

1. *Graduate School of Information Sciences, Tohoku University.*
2. *Research Institute of Electrical Communication, Tohoku University.*
3. *Graduate School of Arts and Letters, Tohoku University*

In order to develop advanced multimedia communications systems, it is important to understand how humans perceive reality from the media presented by the systems. There are various indexes to evaluate the sense of reality. We have been focusing on the index that is the sense of presence and the sense of verisimilitude because we consider that we can apply the figure and ground perception, a knowledge of psychology, to affective perception of space. Furthermore, by adding appropriate sensory information related to the foreground or background component, these perceptual realities could be enhanced. Since sound and vibration have close relationship, the sound of content includes rich information about events. Thus, in this study, whole-body vibration was generated from the sound. We generated nine types of vibration conditions related to the foreground or background component of Audio-Visual contents by adjusting the cutoff frequency and the carrier frequency of the sound. The results showed that higher verisimilitude was observed when vibration closely connected to foreground components was added in a scene. Moreover, under that condition, the sense of presence was hardly affected even when the vibration was added to the content. These results suggest that generating more realism is possible by enhancing vibrations if the processing is performed to match the foreground components in the scene.

Word order and voice influence the time course of sentence processing in Japanese as a second language

Shanglin Xie, Masatoshi Koizumi

Department of Linguistics, Graduate School of Arts and Letters, Tohoku University

Word order in the Japanese language is relatively free. The canonical word order is "subject- object- verb", and the scrambled word order is "object- subject- verb". The "scrambling effect" was observed in previous studies on Japanese sentence processing by native Japanese speakers (Chujo, 1983; Miyamoto & Takahashi 2002; Tamaoka et al. 2005), which revealed that they took longer reading time for processing scrambled word order sentences in comparison with that for canonical word order regardless of whether the voice was active or passive. On the other hand, the Chinese language has the "subject-verb-object" word order, and does not allow any scrambled orders. This difference should cause a difficulty for native Chinese speakers learning Japanese to process sentence scrambling in Japanese. Particularly, the word order of Japanese active sentences (SOV) and Chinese active sentences (SVO) are different, whereas the word order of Japanese passive sentences (SOV) and Chinese passive sentences (SOV) are the same. In order to investigate how native Chinese speaker learning Japanese comprehend Japanese language voice and word order, we used a sentence judgment task for two groups of learners with different proficiency levels. Results revealed voice and word order interactively influenced on sentence processing of Japanese by native Chinese speakers. Their reaction time of the passive sentences was significantly longer than active sentences, and their reaction time of the scrambled word order sentences was significantly longer than canonical word order sentences. The significant effect of word order was weaker in passive sentences than in active sentences. Chinese learners preferred word order that agent precedes patient. In the passive sentences, learners with high proficiency showed a stronger effect of word order than learners with low proficiency, suggesting that learners with higher proficiency process similarly to the native speakers.

How native Japanese listeners normalize the incorrect lexical pitch accent?: An ERP and ERSP evidence

Taiga Naoe[1], Qiong Ma[12], Min Wang[1], Masatoshi Koizumi[1], Shingo Tokimoto[3], & Sachiko Kiyama[1]

[1] *Department of Linguistics, Graduate School of Arts and Letters, Tohoku University*

[2] *Research Institute of National Rehabilitation Center for Persons with Disabilities*

[3] *Faculty of Foreign Language Studies, Mejiro University*

Previous studies suggest that Japanese listeners have the ability to normalize incorrect pitch accent. However, the temporal process of this normalization remains unclear. It has been assumed that there are three stages of spoken word recognition: (1) the pre-lexical stage, which reflects phonological processing; (2) the lexical stage, which activates the mental lexicon and lexical meaning; and (3) the post-lexical stage, which verifies the whole word according to the context. In order to identify at which stage native Japanese listeners normalize incorrect pitch accent, the present study recorded the electroencephalography (EEG) while they engaged in a cross-modal priming task with auditory and visual stimuli. In particular, we utilized indices of event-related potential (ERP) and event-related spectral perturbation (ERSP) to ensure that the processing cost reflected the pitch normalization. First, we investigated ERP for auditorily presented words, to identify when the processing cost for incorrect pitch accent was induced during the word recognition stages. Then, we examined ERSP for the subsequently presented visual stimuli depicting the words with a particular focus on gamma power (30-40Hz), which is known to reflect the integration of intracerebral representation and visual information. The ERP result for auditory stimuli indicated a significant effect of P350, which has been reported to be related to the facilitation/suppression of (2) lexical identification for incorrect pitch accent in comparison with the correct one ($p < .05$). The ERSP indicated that the visual stimuli following word presentation with incorrect pitch accents induced the increased gamma power, in comparison with when the visual and auditory stimuli did not match ($p < .05$), suggesting that words spoken in incorrect pitch accent are normalized to correct accents. The present study concludes that native Japanese listeners normalize a word spoken with incorrect pitch accent at the (2) lexical stage.

Challenge of Energy-efficient AI Hardware Based on Nonvolatile Logic

Takahiro Hanyu[1] and Daisuke Suzuki[2]

[1] *Research Institute of Electrical Communication, Tohoku University,*
  Sendai, Japan 980-8578, hanyu@riec.tohoku.ac.jp

[2] *Frontier Research Institute for Interdisciplinary Sciences, Tohoku University*

Recently, many internet of thing (IoT) applications have arisen in various domains, such as health, transportation, smart homes/cities. In such IoT applications, massive data sets such as yotta-scaled data sets must be processed at each sensor nodes in a real-time fashion. Artificial intelligence (AI) performs a key role for such real-time data processing and there are great demands on low-power, high-performance AI hardware. The fine-grained structure of a field-programmable gate arrays (FPGA) has proven to provide powerful implementations of AI algorithms with better energy efficiency than CPUs and GPUs. However, large standby power consumption is a critical issue for battery-powered or energy-harvested scenarios. A magnetic tunnel junction (MTJ) based nonvolatile FPGA is a promising solution for the standby power problem. In this poster, nonvolatile logic circuit techniques for MTJ-based NV-FPGA and its prospects for an energy-efficient AI hardware are presented.

Lightweight Cryptographic Hardware and Applications in Large Scale Data Encryption for AI

Ville Yli-Mäyry, Rei Ueno, Naofumi Homma
Tohoku University

Interest in Artificial Intelligence has increased in a great number of applications. However, with the adoption of AI applications, the security of the applications has to be taken into consideration. One of the problems for these applications is that providing information security for the handled data is resource intensive. On the other hand, lightweight cryptography has been researched greatly in recent years as the need for encrypted communication and secure memory has grown. In particular, lightweight cryptographic ciphers such as PRINCE offer encryption and decryption at an extremely low latency. In the context of AI and other applications that require large scale data encryption, these low-latency ciphers are attractive for securing the application platform used with full memory encryption, for example. In this study, we present an ultra-low latency implementation of a PRINCE processor, and further propose a countermeasure for improving the tamper-resistance of the design.