The University of Electro-Communications

UEC Research and Innovation

Latest updates on research and innovation at UEC Tokyo.

Vol.1 October 2022

The University of Electro-Communications 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585

Contents

The University of Electro-Communications 1-5-1 Chofugaoka, Chofu, Tokyo 182-8585

News

Delegation of KMITL and CMKL visits UEC

July 12, 2022

https://www.uec.ac.jp/eng/news/announcement/2022/20220712_4635.html

On June 30, 2022, Assoc. Prof. Dr. Anuwat Jangwanitlert, President of King Mongkut's Institute of Technology, Ladkrabang (KMITL), and Assoc. Prof. Dr. Supan Tungjitkusolmun, President of CMKL University (CMKL) visited the University of Electro-Communications (UEC).

UEC President Tano Shunichi, Dr. Abe Koji, Member of the Board of Directors of UEC, Prof. Watanabe Shinichi, Director of International Education Center and other UEC professors welcomed delegation of KMITL and CMKL, and had active discussion for continued and further collaboration between the universities.

KMITL is a university of science and technology established in 1960, and has been producing many excellent researchers since its establishment. KMITL and UEC signed cooperation agreement in 1997 and have been cultivating mutual collaboration through the exchange of faculties and students.

CMKL is established in 2017 in collaboration with KMITL and Carnegie Mellon University, USA. Assoc. Prof. Dr. Supan Tungjitkusolmun, President of CMKL, gave an overview of CMKL and its exchange and internship programs that create international and talented human resources.

We look forward to further development of educational and research exchanges between the universities in the future.

Group photograph

Research Highlights：Research

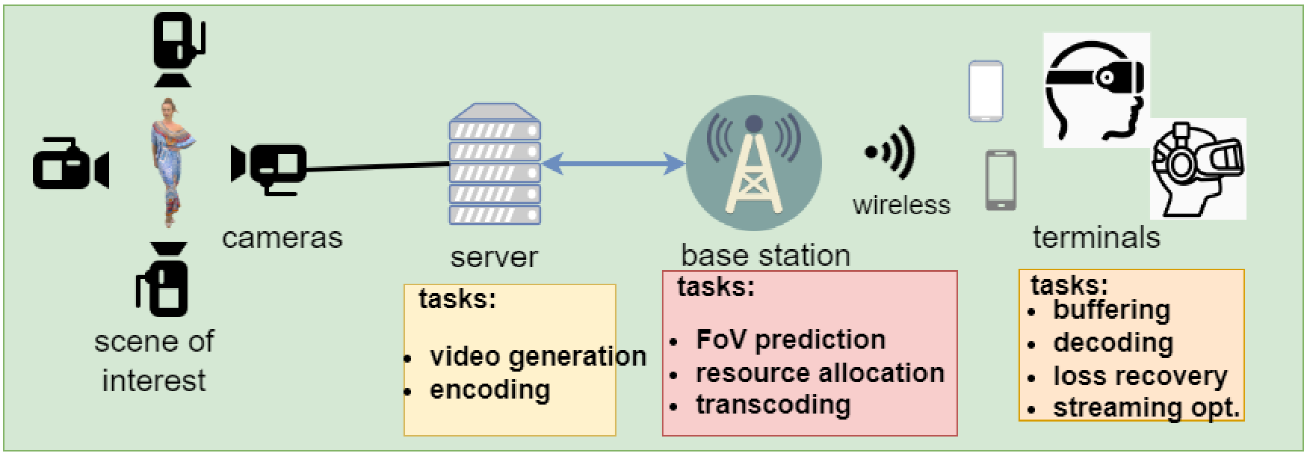
Video technology **Wireless video streaming of the future**

With virtual reality video, users can explore an environment without being physically present. The immersive viewing experience offered by virtual reality is limited, however, because users cannot move forward or backward freely: they can essentially only 'look around' by rotating their head. (Users wear a virtual reality helmet that projects the environment onto their eyes.) This limitation is overcome in volumetric video, also known as hologram video, offering users six degrees of freedom (free forward/backward, up/down, and left/right free motion). Volumetric video can be used for education purposes and in healthcare and is at present being researched intensively. It is expected to become a typical application of 5G (and beyond) wireless communication — indeed, networked volumetric video, allowing users to be untethered, is in critical demand. Still, efficiently transmitting volumetric video data over wireless networks remains a key challenge. Zhi Liu from the University of Electro-Communications, Tokyo, and colleagues have now explored point cloud video streaming, a promising approach for representing and wirelessly transmitting 3D environments.

Point cloud video uses data recorded by several cameras located at different locations and angles around a scene. The scene is divided into volumetric elements called voxels, with each voxel having several attributes such as 3D coordinates and colour. These data form a so-called point cloud. Point cloud video streaming poses several challenges, which essentially come down to the large size of the raw video data, the high complexity of the computations involved, and the high computing power required. Liu and colleagues provide suggestions for overcoming these challenges. For example, rather than transmitting complete video content, it suffices to send only data of the field of view of a user. Also, predicting user behaviour can help to efficiently use available computer resources; such predictions can be based on machine-learning as soon as enough user view logs — needed to train machine-learning models — will become publicly available. In addition, by transmitting uncompressed video tiles (tiles that do not need to be decoded again at the user side but require greater bandwidth) in addition to the compressed video tiles, the computational burden on the user device is greatly reduced.

The researchers demonstrated the feasibility of point cloud video streaming in an actual experiment with 'off-the-shelf' devices (a laptop as the server and a PC as the client). The approach of Liu and colleagues uses a technique called tiling, in which a point cloud video is partitioned into cuboids referred to as tiles. Tiling increases the video streaming efficiency.
Uncompressed video tiles were used to release high computational requirements. The scientists also employed an international-standard encoding scheme called MPEG Dynamic Adaptive Streaming over HTTP (MPEG DASH) and optimized its usage. Comparing original video images to images received by the client showed that the proposed streaming solution provides good quality in an efficient way.

Finally, Liu and colleagues verified the performance of their algorithm in simulations with different numbers of CPU cores and different network bandwidth conditions. More CPU cores and higher bandwidth of course increase the quality of the streaming, but the proposed scheme optimally uses bandwidth and computing resources so that 'video freezes' are fewest and the received video quality is the best. The researchers concluded that they have achieved satisfactory performance and envision "that the point cloud video streaming system will play a vital role in future society" and that it "opens up many exciting and critical future research directions."



Schematic illustration of a typical point cloud video streaming system and the tasks of its components.

Reference

Zhi Liu, Qiyue Li, Xianfu Chen, Celimuge Wu, Susumu Ishihara, Jie Li and Yusheng Ji, Point Cloud Video Streaming: Challenges and Solutions, *IEEE Network* **35**, 202-209 (2021).

URL: https://ieeexplore.ieee.org/document/9537928
DOI: 10.1109/MNET.101.2000364

Zhi Liu, Associate Professor, Department of Computer and Network Engineering

Website

https://liuzhiabc.github.io/

Video Profiles：Research

Perceiving the world through tactile perception

Human tactile perception
Hiroyuki Kajimoto
Professor, Graduate School of Informatics and Engineering

Hiroyuki Kajimoto is conducting research on human tactile perception, that is, how people perceive tactile sensation, why the illusion phenomenon occurs, and how to use tactile sensation for virtual reality (VR) applications.

“We are currently experiencing serious consequences related to COVID-19 and other issues,” says Kajimoto. “Over the past two years many people were not able to go to school or their office. We have realized that our freedom to move is fragile and that can be taken away suddenly. On the other hand, we have discovered that information technology, such as online conferencing, is a powerful tool for overcoming restrictions on movement. But I believe that there is still something missing from conventional technologies, specially, the sense of touch. So, in my laboratory we are conducting research on how to artificially produce a realistic sense of touch, for the VR and meta-verse world.”

Perceiving the world through tactile perception

In this research, the first major research topic is to clarify how people perceive the world through tactile perception, for example, to observe physical phenomenon, and elucidate the phenomenon of tactile illusions.

Kajimoto explains that when we touch an object, our skin makes contact with it and it is clear how the skin deforms under such circumstances. But it is difficult to observe deformation of the skin using a camera because the actual object has a rough surface. Kajimoto and colleagues developed a method in which the finger and objects being touched were all submerged in oil to make them optically transparent to enable the first ever successful observation of the deformation of a finger in contact with an object.

The sense of touch encompasses many illusory phenomena. For example, when two hands continue touch each other, it feels as though that there is a thin plate in between the hands. This illusion indicates that human perception of ‘flatness’ is ambiguous. Kajimoto hypothesized that this phenomenon is caused by symmetrical deformation of fingers and verified his hypothesis through experiments.

“Some tactile illusions produce a sense of force,” explains Kajimoto. “We have named one of the most powerful of such forces as the ‘hanger reflex’ referring to the case when a hanger is placed on the head, and the head rotates to the left or right. This phenomenon indicates that cutaneous sensation can be misinterpreted as a force sensation, and we move according to the illusory force. We have identified the locations involved in this phenomenon and confirmed that a similar illusion can be produced by something as simple as a rubber band.”

Presenting realistic tactile sensations for virtual reality and remote communications

The second research topic being explored by Kajimoto is "how to present tactile sensations to people”, for virtual reality and remote communication applications.

The most compact method of presenting tactile sensations is by electrical stimulation, where stimulating nerves inside the skin enables the presentation of distinct tactile sensations. “We have realized a tactile presentation device with more than 1,500 stimulation points, transparent electrodes on a smartphone, and high-density tactile display for fingertips,” says Kajimoto. “We are also developing new methods of tactile presentation. For example, force sensation is produced by electrical stimulation of the skin, cold temperature sensation is produced by electrical stimulation, we have designed a new vibration transducer, and temperature sensation was produced using chemical substances.”

An example of a successful link between the study of illusion and its application is the study of the hanger reflex. Research on the hanger reflex was initially conducted on the head, but it is now clear that the same phenomenon occurs in other parts of the body, such as the waist. Kajimoto and colleagues have this phenomenon to induce movement when walking or riding a vehicle such as a Segway.

In most cases, tactile sensation is presented to the fingertips for VR applications. However, Kajimoto has presented tactile sensation to the forearm or to the face with the goal to create a tactile device that can be worn more easily and still have a certain realistic and immersive feeling.

This research is an attempt to utilize a classically known illusion called kinesthetic illusion. It is known that vibration of tendons and muscles creates the illusion that the arm is moving. By utilizing this phenomenon, it is possible to present the sensation that the body is tilting, or the weight of a stick is changing in VR environments.

“My laboratory conducts research in science, engineering, and applications, mainly in tactile sensation. We are actively accepting students from overseas, so please consider joining us.”

References and further information

1. Jianyao Zhang, Hiroyuki Kajimoto, Approaches for Reproducing the Haptic Sensation of Sandpaper with Different Roughness During Bare Fingertip Interaction, Frontiers in Virtual Reality 3/ 1, 2673-4192 (2022).
2. Taha Moriyama, Hiroyuki Kajimoto, Wearable Haptic Device Presenting Sensations of Fingertips to the Forearm, IEEE Transaction on Haptics 15/ 1, 91-96, (2021).
3. Taha Moriyama, Hiroyuki Kajimoto, HARVEST: High-Resolution Haptic Vest and Fingertip Sensing Glove That Transfers Tactile Sensation of Fingers to the Back Joint, Applied Science, 11/ 1298, (2021).

Website
https://kaji-lab.jp/ja/index.php

Video Profiles：Innovation

Unsupervised segmentation and its applications

Unsupervised segmentation deep learning
Tomoaki Nakamura
Associate Professor, Graduate School of Informatics and Engineering

Tomoaki Nakamura conducts research on unsupervised segmentation of time-series data based on hierarchical Bayesian models.

Unsupervised segmentation and its applications

“Unsupervised means that a person does not have to give the correct answer. In recent machine learning, supervised learning—which learns from data with correct labels attached in advance—is mainstream,” explains Nakamura. “On the other hand, my approach has the advantage of being able to handle data without the correct labels. Next, segmentation technology extracts meaningful patterns from time-series data and classifies them into groups of meaning. An example is spoken language, where sound is in the form of vibrations of air. Humans divide it into phonemes, which are units of sound, and recognize them discretely. Furthermore, each phoneme sequence is discretely recognized by dividing it into words, which are units with meaning. In this way, segmentation is a technique for segmenting time-series data, grouping it according to meaning, and handling it discretely. Unsupervised segmentation enables analysis of unlabeled time series data.”

Focus on generative models for unsupervised segmentation

One unsupervised segmentation method is Gaussian Process-Hidden Semi-Markov Models (GP-HSMM). This model assumes that time-series data is generated by creating segments, which are partial time series, from the association patterns of multiple classes and connecting them. A model that expresses the process by which data is generated in this way is a called a generative model. In a generative model, by probabilistically describing the process of data generation, unsupervised learning that probabilistically infers parameters from data alone is possible.

The purpose of GP-HSMM training is to estimate parameters that maximize the probability of segment ‘X’ being generated from time-series data ‘S’ only. In other words, what kind of fixed patterns are there from data alone?　From which fixed pattern are the partial time series in the data generated? And how long is that segment that must be estimated? Nakamura and his group have developed a method for efficiently inferring these parameters.

Here are two examples of actual analysis cases using this method. The first is human motion analysis, where the researchers extracted basic motions included in human motion from only 96-dimensional motion capture data of human motion.

In the accompanying video, the graph on the left is the input time-series data, and the graph on the lower right is the result of segmenting this data only. The horizontal axis represents time, and the vertical axis represents the number of basic movements classified. In other words, sections assigned the same number represent the same basic movement. If this is compared with the motion capture data visualized in the middle video, the class changes when the action changes. In this way, it is possible to extract six basic motions, which are automatically put into meaningful groups, from only the time-series data.

Motion capture data of a marmoset

Another example is related to animals where this method was successfully applied to the motion capture data of a marmoset, rapidly segmenting it, and automatically extracting characteristic behavior. Furthermore, Nakamura extended this method, and developed a model to segment data with a double segmentation structure. For example, speech described above becomes data with a double segment structure.

“By segmenting the speech waveform, we can extract phonemes, and by segmenting the phoneme sequence, we can extract words,” explains Nakamura. “To extract words from speech waveforms, two stages of segmentation are required. Therefore, by hierarchically connecting segmentation models, we applied this method to speech with a double segmentation structure.”

Notably, in this model a segmented model called HSMM is placed on top of the GP-HSMM model described earlier. In this model, all characters, sound characteristics, and words can be learned without supervision from speech waveforms alone.

The video shows the results of the actual analysis of voice using a data set of artificial words composed of the sounds of "aiueo". As words, it consists of two-word sentences such as “a-oi” and “ao” and three-word utterances.

In the video, the graph on the left is the result of segmenting only the speech waveform, and the unit "AIUEO" was correctly segmented as a character. Furthermore, by segmenting this syllable string the words contained in the dataset were extracted. The graph on the right is an index representing the closeness to the correct answer called ARI. This method achieves highly accurate segmentation compared to other methods.

Future

"In this video I introduced a technique for unsupervised segmentation that can extract fixed patterns from time-series data without the correct labels," says Nakamura. "I introduced GP-HSMM as a basic model and demonstrated that it is possible to extract basic motion and characteristic motion from motion capture data. In addition, I introduced a double segmentation analysis model that hierarchically connects segmentation models and showed that syllables and words can be extracted from speech alone. In the example introduced this time, only motion and voice were used, but this method can also analyze various time-series data, that are not limited to motion and voice."

Research Keywords: motion segmentation, Gaussian process, hidden semi-Markov model, motion capture data, high-dimensional time-series data

References and further information

1. Masatoshi Nagano, Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi and Wataru Takano, "HVGH: Unsupervised Segmentation for High-dimensional Time Series Using Deep Neural Compression and Statistical Generative Model", Frontiers in Robotics and AI, Vol. 6, Article 115, pp. 1-15, Nov. 2019

2. Masatoshi Nagano, Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi and Wataru Takano, "High-dimensional Motion Segmentation by Variational Autoencoder and Gaussian Processes", IROS2019, pp. 105-111, Nov. 2019

3. Masatoshi Nagano, Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi, Masahide Kaneko, "Sequence Pattern Extraction by Segmenting Time Series Data Using GP-HSMM with Hierarchical Dirichlet Process", IROS2018, pp. 4067-4074, Oct. 2018

4. Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi, Hideki Asoh and Masahide Kaneko, "Segmenting Continuous Motions with Hidden Semi-Markov Models and Gaussian Processes", Frontiers in Neurorobotics, vol.11, article 67, pp. 1-11, Dec. 2017

Patent submissions (Japanese)

1. 中村友昭，長井隆行，岩田健輔，情報処理装置、情報処理方法およびプログラム，特許第 6908286 号（2021/07/05），PCT/JP2017/032109（出願日 2017/9/6），WO2018/047863（公開日 2018/03/15），特願 2016-174085（出願日 2016/9/6）

2. 中村友昭，長井隆行，池田佳那，情報処理装置、情報処理方法およびプログラム，特許第 6829368 号（2021/01/26），特願 2016-235922（出願日 2016/12/5）

Website

https://hp.naka-lab.org/