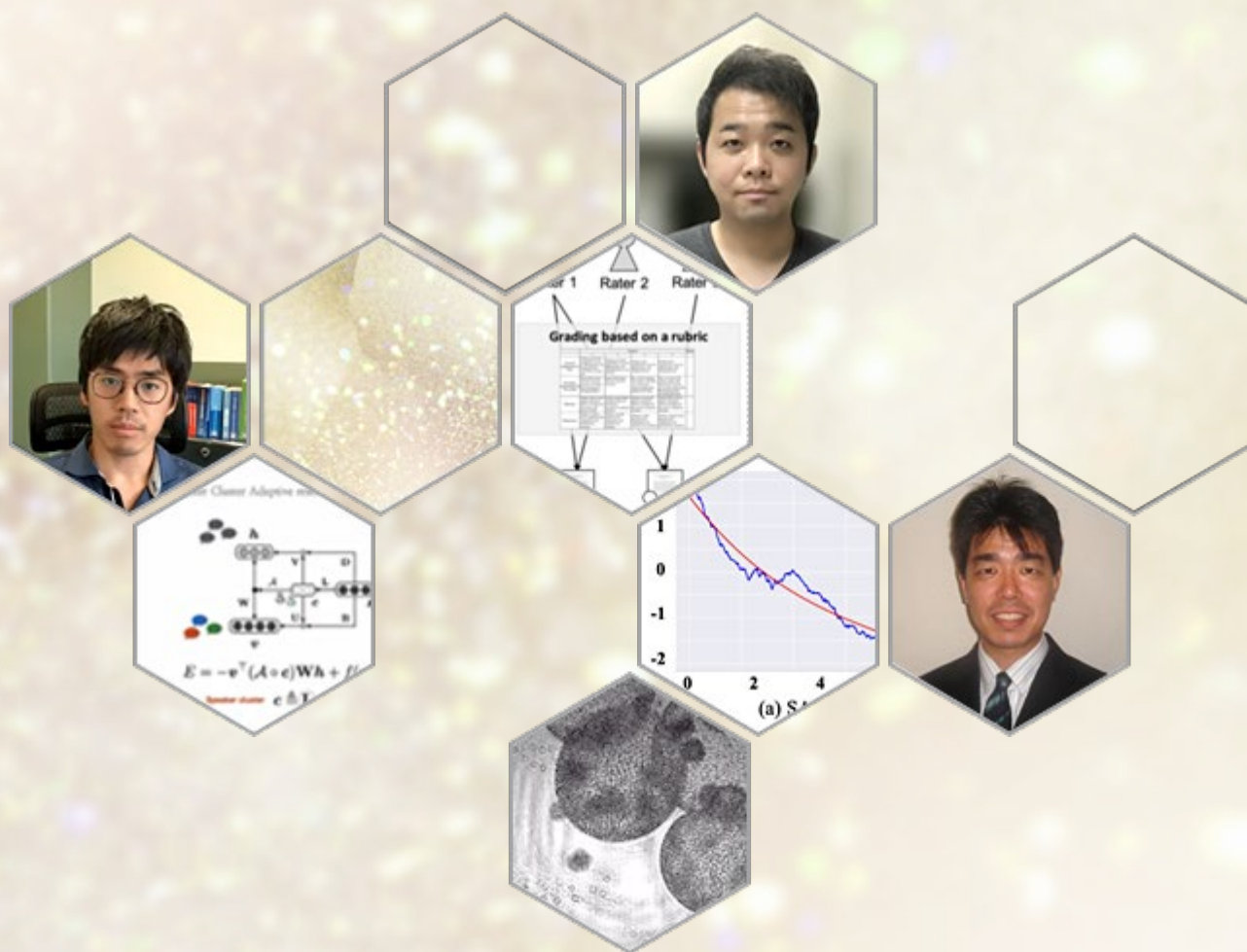**UEC** The University of Electro-Communications

UEC e-Bulletin

Updates on research, innovation, and events at UEC:
Unique and Exciting Campus in Tokyo

Vol.30, December 2021

Research Highlights

- Frequency analysis helps to understand sleep disorder

- Modelling performance assessment

Researcher Video Profiles

- Speech Signal Processing Based on Shallow Neural Networks

Topics

- Fascination with digital holograms and their applications for imaging through semi-opaque materials

Sleep science

Frequency analysis helps to understand sleep disorder

Sleep apnea syndrome (SAS) is a sleep disorder characterized by the occurrence of pauses in breathing (apnea) during sleep.   Such pauses can typically last for more than 10 seconds and are often followed by loud snoring.   The brain interprets each breathing pause as danger — because of the decrease in oxygen supply — and sleep becomes shallow.   As a result, a person suffering from SAS builds up a sleep debt, which may in turn lead to mental health issues like depression or dementia.   In order to avoid medical complications, early detection of SAS is crucial.   So-called non-contact detection methods are based on monitoring chest motion, e.g. by means of a sensor attached to the mattress sensor the person is sleeping on; from the recorded bio-vibration data, breathing frequencies and amplitudes can be derived.   This type of method is not always effective.   For example, when a person's breathing is 'forced' (breathing accompanied by thoracic and abdomen movement, and in fact also a symptom of SAS), sleep apnea is difficult to detect.   Now, Iko Nakari and Keiki Takadama from the University of Electro-Communications have developed a new method for processing overnight bio-vibration data that can detect SAS in a more universal way.

The researchers analysed bio-vibration data recorded from 9 SAS patients and 9 healthy individuals, obtained by means of a mattress sensor.   Rather than looking only at respiration (between 0.1 Hz and 0.2 Hz) and heartbeat (between 0.6 Hz and 1.5 Hz) frequencies, they considered frequencies up to 8 Hz, and looked at the distribution — the spectrum — of frequencies.   When comparing frequency spectra, Nakari and Takadama noticed a slight increase in frequency density around 3 Hz for the SAS patients.   On a logarithmic plot of the frequency spectrum, this increase manifests itself as a convex shape.   Based on this observation, the researchers defined a quantity called the degree of convexity of the logarithmic spectrum (DCLS).

Remarkably, the average DCLS value for the SAS patients ($\approx 99 \pm 10$) is completely separate from the average value for the healthy subjects ($\approx 48 \pm 7$).   Therefore, the DCLS value has the potential to be used as an indicator for SAS — obtained just by sleeping on a mattress sensor.

Further analysis showed that the increased frequency density around 3 Hz corresponds to accumulated density in the so-called WAKE stage (the first of six levels used for characterizing 'sleep deepness').   Therefore, it is likely that the WAKE stage is different for SAS patients and people not suffering from sleep apnea.   Even more, the researchers argue that SAS subjects generate 3 Hz waves during WAKE phases, and believe that this may actually be a hitherto unknown symptom of SAS, apart from the apnea itself.   However, as Nakari and Takadama point out, future work "should clarify the phenomenon around 3 Hz".
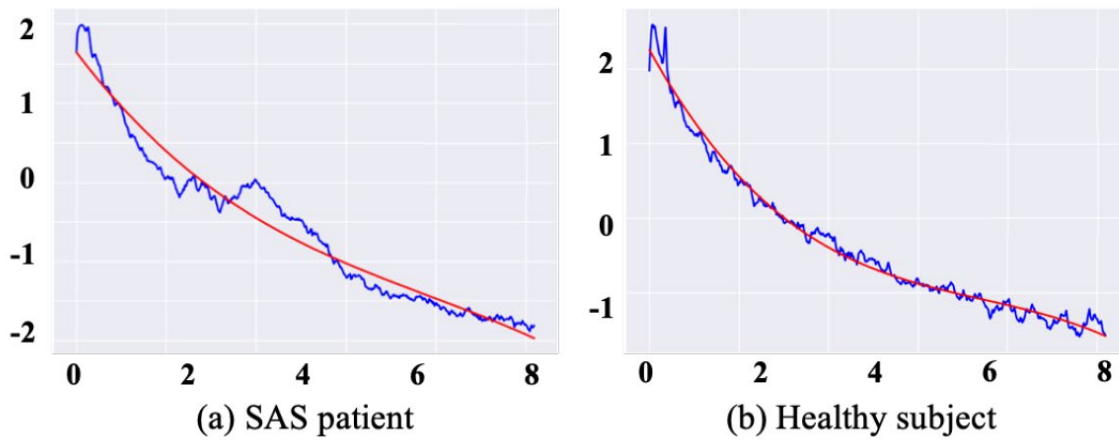
Figure: Exemplary averaged logarithmic spectrum of a SAS patient and a healthy person.

Iko Nakari and Keiki Takadama, Sleep Apnea Syndrome Detection Based on Degree of Convexity of Logarithmic Spectrum Calculated from Overnight Bio-vibration Data of Mattress Sensor, pp. 2274–2277, (2021).

The 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC2021) (2021).

URL: https://embc.embs.org/2021/

Educational measurement

Modelling performance assessment

Performance assessment of a practical task carried out by an examinee is typically done by human raters awarding scores for different parts of the task. Often, a so-called scoring rubric is used for this purpose, listing the various parts and descriptions of the performance scores associated with them. There are some inherent shortcomings to this procedure, however, including the characteristics of the rubric's evaluation items and the raters' behaviour — one rater may score differently than another. Now, Masaki Uto from the University of Electro-Communications has developed a new model that takes into account the specifics of a rubric's evaluation items and the raters.

The approach followed by Uto relies on models developed in a theoretical framework known as item response theory. It is based on a formula giving the probability P_{ijkr} that examinee j gets score k for evaluation item i by rater r. The formula typically contains parameters such as the difficulty (β_i) for the evaluation item, the latent ability of the examinee (θ_j) and the severity of the rater (β_r). The idea is then that, by fitting the formula to an existing dataset with known score outcomes, good values of the parameters (like β_i, θ_j and β_r) can be obtained. Yet, this description is almost always too simplistic to result in good results, however.

One improvement lies in incorporating the notion of ability dimensions — an abstract representation of an examinee having different ability 'spheres'. Uto's model combines ability dimensions with rater characteristics, which signifies a step forward in item response theory modelling.

Apart from providing a more realistic description of performance assessment with a rubric and raters, the model can also help to check the quality of the rubric's evaluation items, as well as providing insights into what exactly each ability dimension measures.

Uto tested the probability formula by first simulating a large number of data sets, with randomly generated parameters. Then, the data sets were fitted to the formula, resulting in estimated parameters. Good agreement between the true and the fitted parameters was obtained, showing that the model works well. Moreover, specific simulations showed that the inclusion of rater characteristics led to improved examinee ability accuracy.

The model was also tested in actual data experiments, with 134 Japanese university students performing an essay-writing task requiring no preliminary knowledge. One conclusion was that, for this case, a two-dimensionality assumption worked better than a one-dimensional ability. A further finding was that the inclusion of rater characteristics indeed improved model fitting.

Uto plans to further test the model's effectiveness using various and more massive datasets, and to, quoting the researcher, "extend the proposed model to four-way data consisting of examinees × raters × evaluation items × performance tasks because practical tests often include several tasks."
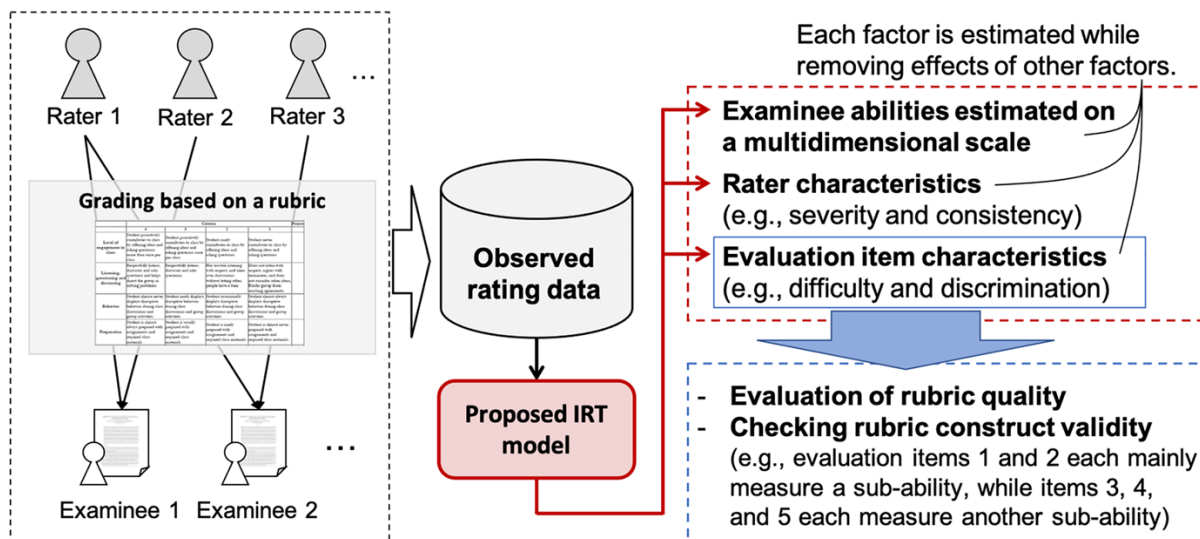

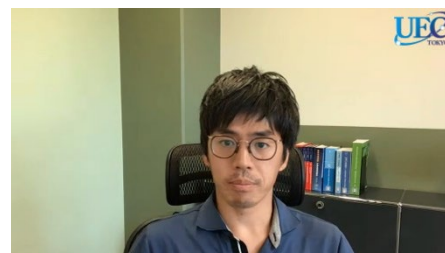Figure: Proposed model to analyse rubric-based scoring.

Speech Signal Processing Based on Shallow Neural Networks

Toru Nakashika, Associate Professor
Department of Computer and Network Engineering
Graduate School of Informatics and Engineering
University of Electro-Communications, Tokyo

In this video feature Toru Nakashika describes his group's research on speech signal processing using shallow neural networks.

It is widely known that deep learning (DL) is used in audio signal processing. In this approach, many studies use DL by increasing the number of layers of neural networks and parameters in the "dark cloud" to improve expressiveness and accuracy.

However, such models have problems such as high calculation costs and the need for huge amounts of data. Furthermore, DL is often called a black box, and it is difficult to interpret what is being done internally. Therefore, it is difficult to come up with ideas for improvements.

"The goal of my research is to produce the same level of accuracy in speech recognition and synthesis as in deep learning but by using interpretable and shallow models based on appropriately expressing the structure of speech data," explains Nakashika. "That is by using wisdom instead of computational resources, we aim to reduce computational costs and achieve more practical speech recognition and speech synthesis."

Nakashika and his colleagues use shallow models, including the Boltzmann machine model—an example of a shallow and interpretable model. The use of a Boltzmann machine enables the expression of an arbitrary probability distribution by freely designing so-called called energy functions, and audio data structures can be appropriately expressed using this model.

Since this Boltzmann machine is a shallow model, it has the advantage that both calculation costs and the amount of data required for learning reduced can be significantly reduced.

Some recent results obtained by Nakashika include voice identity conversion—a technology that processes voice and only converts a person's personality without changing the contents of the utterance. "I have proposed a model called the speaker-cluster-adaptive restricted Boltzmann machine," says Nakashika. "This is an extension of the Boltzmann machine, and conversion is possible with only about one second of data."

Nakashika has also proposed the so-called complex-valued restricted Boltzmann machine model that directly expresses complex numbers. Sound is often expressed in a complex spectrum, but since it is known that amplitude is better recognized by humans than phase, it is possible to omit the phase and only use the amplitude spectrum. "I think that it would be more expressive if there was a model that could directly express the phase, and the model that can directly express the complex spectrum of the voice is the complex-restricted Boltzmann machine mentioned earlier," says Nakashika. "We showed that this makes it possible to synthesize speech with higher accuracy than the conventional VOice enCODER."

Plans include the application of the Boltzmann machine to speech synthesis and voice quality conversion in other fields of speech signal processing, such as speech recognition and sound source separation. "I would like to encourage more promote more research on shallow neural networks."

References and further information

・Toru Nakashika, and Kohei Yatabe, "Gamma Boltzmann Machine for Audio Modeling,"IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol.29, pp.2591-2605, July 2021.
DOI:10.1109/TASLP.2021.3095656

・Takuya Kishida and Toru Nakashika, "Speech chain VC: linking linguistic and acoustic levels via latent distinctive features for RBM-based voice conversion," IEICE TRANSACTIONS on Information and Systems, Vol.E103-D, No.11, pp.1-11, August 2020.
DOI:10.1587/transinf.2020EDP7032

・Kentaro Sone and Toru Nakashika, "Pre-Training of DNN-Based Speech Synthesis Based on Bidirectional Conversion between Text and Speech," IEICE TRANSACTIONS on Information and Systems, Vol.E102-D, No.8, pp.1546-1553, August 2019.
DOI:10.1587/transinf.2018EDP7344

・Toru Nakashika, Shinji Takaki, and Junichi Yamagishi, "Complex-Valued Restricted Boltzmann Machine for Speaker-Dependent Speech Parameterization From Complex Spectra,"IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol.27, No.2, pp.244-254, Oct. 2018.
DOI:10.1109/TASLP.2018.2877465

・Toru Nakashika and Yasuhiro Minami, "Speaker-adaptive-trainable Boltzmann machine and its application to non-parallel voice conversion," EURASIP Journal on Audio, Speech, and Music Processing, DOI: 10.1186/s13636-017-0112-6, pp.1-10, June 2017.
DOI:10.1186/s13636-017-0112-6

・Toru Nakashika, "Deep Relational Model: A Joint Probabilistic Model with a Hierarchical Structure for Bidirectional Estimation of Image and Labels," IEICE Transactions on Information and Systems, Vol.E101-D, No.2, pp.428-436, Feb. 2018.
DOI:10.1587/transinf.2017EDP7149

・Toru Nakashika, Tetsuya Takiguchi, and Yasuhiro Minami," Non-Parallel Training in Voice Conversion Using an Adaptive Restricted Boltzmann Machine," IEEE/ACM Transactions on Audio, Speech and Language Processing, Vol.24, No.11, pp.2032-2045, Nov. 2016.
DOI: 10.1109/TASLP.2016.2593263

Website

http://www.sp.lab.uec.ac.jp

Fascination with digital holograms and their applications for imaging through semi-opaque materials

Eriko Watanabe, Associate Professor, Department of Engineering Science, UEC Tokyo

"My interest in light and optics was triggered when I saw an exhibition on holography at an event on campus during my undergraduate days at university," says Eriko Watanabe, an associate professor at the Department of Engineering Science, UEC Tokyo. "I was intrigued by the amazing three dimensional optical structures that could be produced by the interference of light waves. This fascination with holograms is the basis for my current research."

Recent research being conducted by members of the Watanabe Group includes digital holography imaging of objects hidden by media such as scatter plates and biological tissues. "The ultimate goal is to develop technology for the non-invasive imaging of living cells inside biological tissues," explains Watanabe. "We expect our research will play an important role in clarifying biological mechanisms governing human health on the cellular level." Other potential applications of this technology include imaging through fog and air turbulence, where the latter is important for land-based astronomy where movements of the air can adversely affect astrophotography.

Specific scientific issues to resolve to achieve these goals are (1) elimination of temporally fluctuating spatial noise due to complex fluctuations and scatterers to capture images behind obtrusive objects, and (2) development of microscopic imaging technology for visualizing below living skin.

One solution proposed by the Watanabe Group is using deep neural networks to suppress temporally fluctuating spatial noise and applying optical correlation imaging. "Our imaging method combines deep learning with optical correlation imaging that accelerates ordinary single pixel imaging by the use of optical computing," explains Watanabe. "Furthermore, we are imaging behind scattering media by phase shift digital holography using near-point light sources with planar waveguides. Using a near-point light source eliminates fluctuations with common optical path digital holography and planar waveguides take us closer towards 'needle-type' probe structures."

Recent results

● **Three-dimensional imaging through thick phase-fluctuating media [1]**

Recently, Manami Ohta reported on the experimental and theoretical confirmation of suppression of three-dimensional phase fluctuations using phase-shift digital holography with two adjacent light sources for common-path geometry. The experiments employed Toppan Test Charts with three pieces of glass coated in transparent paint acting as the phase fluctuating layers. The experimental results showed the reconstructed images to be much higher quality than those obtained by conventional approaches. Notably, Manami Ohta received the Student Paper Award for this contribution.

● **Development of innovative high-speed color digital holographic microscope technology [2]**
In this report Kazutaka Nakama—who received a Student Award for this presentation—describes the latest version of compact and easy to use digital holographic microscopes being developed by the Watanabe Group. Specifically, Nakama explains the design of a color planar light wave circuit digital holographic microscope (PLC-DHM) that was successfully used for 3D imaging and 'quantitative phase video imaging at multiple wavelengths."

● **Single-pixel imaging with deep neural networks [3]**
Kazumasa Kimura was awarded the Best Paper Award for this contribution on the development of an effective 'coding pattern' based on deep neural networks for capturing low noise 3D parallax images with a combination of single pixel imaging and integral photography. This approach significantly reduces random noise that is observed using conventional methods.

Contributing to society

"I worked on the development of holographic memories during the early days of my career," says Watanabe. "The applications included storing music and video for the entertainment business. During this research I realized that the digital entertainment sector was being overwhelmed with pirated versions of songs and videos; clear copyright infringement. So I decided to partner with prominent groups in the entertainment business to bring these issues to the notice of the general public by organizing UEC accredited courses on these topics. It's been a highly successful way of educating people about copyright issues with the goal of eliminating such behavior." Examples of Watanabe's activities include collaborating with the Music Publishers Association of Japan and Recording Industry Association of Japan on arranging a series of special lectures themes such as 'Entertainment Business and Copyright in the era of AI'; and 'Game Industry'.

Syllabus of the course on "Special Lecture on Informatics and Engineering B : Copyright and creative business (In Japanese only)

http://kyoumu.office.uec.ac.jp/syllabus/2017/31/31_21017254.html

References

1. M. Ohta, S. Kodama, Y. Miyamoto, W. Osten, M. Takeda and E. Watanabe. Three dimensional imaging through thick phase-fluctuating medium based on phase-shift digital holography with two adjacent light sources for common-path geometry, OSA
Imaging and Applied Optics Congress (Digital Holography and Three-Dimensional Imaging), DM5E.2 online, 2021.7.19 , OSA Student Paper Award Finalist, 2021

2. K. Nakama, H. Gomi, K. Hayashi, K. Okamoto, E. Watanabe. High-speed color digital holographic microscope besed on a planar lightwave circuit with a thin film heater, Imaging, Sensing, and Optical Memory 2021 (ISOM2021), Tu-C-03 online, 2021.10.5 , Student Award

3. K. Kimura, M. Nagata, Y. Wada, Y. Katano, N. Ishii, E. Watanabe, and T. Muroi, Reduction of Random Noise in Parallax Images Acquired by Single-Pixel Imaging using Deep Neural Network, Imaging, Sensing, and Optical Memory 2021 (ISOM2021), Mo-E-04 online, 2021.10.4, Best Paper Award.
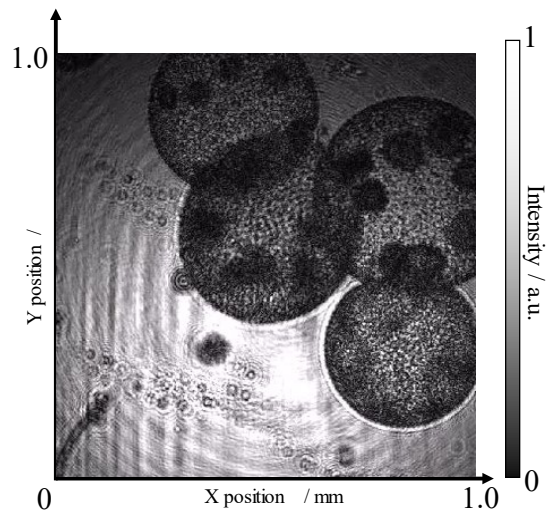
Further information

Eriko Watanabe Group Website

https://sites.google.com/gl.cc.uec.ac.jp/erikowatanabe
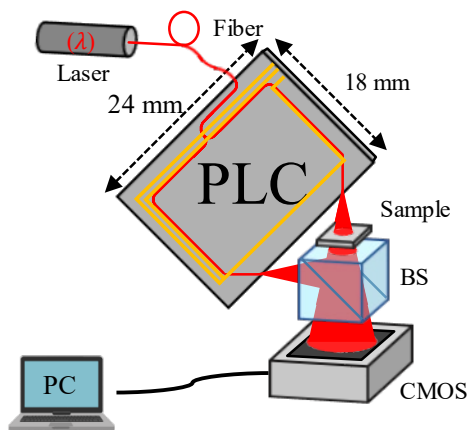
high-speed digital holographic microscope
using planar lightwave circuit(PLC)

Reconstructed image of volvox

The University of Electro-Communications (UEC) in Tokyo

is a small, luminous university at the forefront of pure and applied sciences, engineering, and technology research. Its roots go back to the Technical Institute for Wireless Commutations, which was established in 1918 by the Wireless Association to train so-called wireless engineers in maritime communications in response to the Titanic disaster in 1912. In 1949, the UEC was established as a national university by the Japanese Ministry of Education, and moved in 1957 from Meguro to its current Chofu campus Tokyo.

With approximately 4,000 students and 350 faculty, UEC is regarded as a small university, but with particular expertise in wireless communications, laser science, robotics, informatics, and material science, to name just a few areas of research.

The UEC was selected for the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Program for Promoting the Enhancement of Research Universities as a result of its strengths in three main areas: optics and photonics research, where we are number one for the number of joint publications with foreign researchers; wireless communications, which reflects our roots; and materials-based research, particularly on fuel cells.

International Public Relations

The University of Electro-Communications

1-5-1 Chofugaoka, Chofu, Tokyo 182-8585

E- mail : kokusai-k@office.uec.ac.jp

Website: http://www.uec.ac.jp/eng/