ABSTRACTS

Plenary Speech 1 (1F Conference Room)

January 21 (Wednesday), 11:00-11:50

Chair: Kazushi Ikeda (Nara Institute of Science and Technology, Japan)

Statistical Mechanics Meets Microbial Data: Revealing Network Structure and Chaos in Evolutionary Dynamics

Kei Tokita

Department of Complex Systems Science,
Graduate School of Informatics, Nagoya University

In this talk, I will explore how statistical mechanics, when integrated with microbial time-series data, can reveal the hidden structure and complex dynamics of evolutionary processes in microbial communities. I will begin by presenting key theoretical results derived from statistical physics approaches to evolving complex systems. This will be followed by a case study in which we inferred interaction parameters of a Lotka–Volterra-type model from time-series data of operational taxonomic unit (OTU) abundances in the mouse gut microbiome. The inferred interaction networkexhibited characteristic structural features, and long-term simulations based on the inferred parameters showed quasi-periodic chaotic behavior—suggesting the prevalence of high-dimensional chaos in natural microbial ecosystems. Further, simulations using generalized evolutionary models based on the same parameters produced more realistic OTU dynamics, bridging theoretical modeling with empirical observations. I will also present recent findings from our participation in the Moonshot Research and Development Program, Cabinet Office, Japan, which targets an 80% reduction in methane emissions from cattle by 2050. Our analysis of OTU-level rumen microbiota data has provided new insights into the ecological and evolutionary underpinnings of microbial methane production. Overall, this talk demonstrates how the interplay between statistical mechanics and microbial data can illuminate the network structure and chaotic dynamics that drive microbial evolutionary processes.

Biography:

Kei Tokita is a Professor in the Department of Complex Systems Science at the Graduate School of Informatics, Nagoya University, where he also serves as Vice Dean since 2023. He received his Dr. Sci. in theoretical physics from the University of Tokyo in 1994. He began his academic career as an Assistant Professor in the Department of Physics at Osaka University (1994–2000), and later served as Associate Professor at the Cybermedia Center (now the D3 Center) at Osaka University (2000–2013). He was a Visiting Scholar at Harvard University twice—first in the Department of Biology and Chemistry (Shakhnovich Lab, 1996–1998) and later at the Program for Evolutionary Dynamics (Director: Prof. Martin Nowak, 2003–2004). His research centers on mathematical modeling of complex systems, with a particular focus on ecological and evolutionary dynamics. He has published extensively on community ecology, microbial population dynamics, and evolutionary game theory, using tools from nonlinear dynamics, statistical physics, and network theory. From 2023 to 2024, he served as President of the Japanese Society for Mathematical Biology. His recent work explores chaotic behavior in microbial ecosystems and the controllability of ecological networks using data-driven approaches. With an academic background that bridges physics, biology, and informatics, he seeks to uncover universal principles that govern the organization and dynamics of complex biological systems.

Plenary Speech 2 (1F Conference Room)

January 21 (Wednesday), 15:00-15:50

Chair: Reiji Suzuki (Nagoya University, Japan)

Plural Intelligences: Tacit Knowledge Across Fragmented Systems

Olaf Witkowski

Cross Labs

Complex intelligences—biological, artificial, and hybrid ensembles—can respond to stress by fragmenting control among plural agents. While this adaptive architecture successfully distributes functions and responsibility, it can also complicate coordinated action planning, negotiation of control, integration of partial experiences, and the maintenance of coherent problem-solving strategies. The most challenging—yet crucial—information to share among agents is tacit knowledge (Polanyi, 1958): inherently embodied, context-dependent, experiential insights not easily codifiable into explicit representations. This talk explores how cross-substrate communication—translating internal knowledge across diverse contexts, physical constraints, and embodiments—can enable fragmented systems to coordinate, and sustain high levels of problem-solving. Drawing on examples from robot collectives, cognitive-linguistic effects, technobiological hybrids, but also human minds under chronic stress, and brain–machine study cases, we consider how plural architectures can inform advances in cognition-assistive technologies. We also anticipate what may come after superintelligences: can we fluidly reconfigure identity while preserving coherence, with the “driver” dynamically shifting between machines, humans, and hybrid constellations of diverse other beings?

Biography:

Olaf Witkowski is a pioneering leader in exploring the language of artificial minds. He serves as Founding Director of Cross Labs, an innovative AI research institute in Kyoto dedicated to diverse intelligences and compassionate technology. He currently holds multiple influential positions, including President of the International Society for Artificial Life (ISAL) and Board Director at leading AI company Cross Compass Ltd. He holds a PhD in Computer Science from the University of Tokyo, is an alumnus member of Princeton’s Institute for Advanced Study, and serves as external faculty at various universities in Japan. He has co-founded research ventures across three continents, including YHouse Inc. in New York—a nonprofit focused on consciousness and AI—the Center for the Study of Apparent Selves in Kathmandu, which explores Buddhist philosophy and AI ethics, and the newly established Artificial Life Institute in Kyoto.

Plenary Speech 3 (1F Conference Room)

January 22 (Thursday), 11:00-11:50

Chair: Fumitoshi Matsuno (Osaka Institute of Technology / Kyoto University, Japan)

The Next Standard in Robotic Surgery with AI: Flexible Endoscopic Surgery Robot "Zamenix"

Dong-Soo Kwon

Founder & CEO, ROEN Surgical Inc. / Emeritus Professor, KAIST

The surgical robot system, such as da Vinci, has changed the paradigm of laparoscopic surgery by proving the value of robotic precision, dexterity, and ergonomics. Today, the next paradigm shift is emerging—minimum-invasive and non-invasive endoluminal surgery enabled by flexible endoscopic robot systems. **Zamenix** is a robotic endoluminal surgery system for kidney stone treatment that teleoperates a flexible ureteroscope, a laser fiber, and a stone basket. Zamenix enables precise, stable teleoperation of an endoscope and surgical instruments, offering functions for laser fragmentation/dusting of stones and retrieval/suction of fragmented stones. It provides intuitive remote manipulation without the need for an X-ray protection gown. It includes features such as automation of repetitive tasks, detection of oversized stones, and respiratory-motion compensation via AI technology. A first-in-human trial in Korea demonstrated favorable stone-free rates with minimal complications. Since its launch in 2024, 15 units have been installed in leading university hospitals, and more than 300 clinical cases have shown the efficacy and efficiency of robotic surgery. During the commercialization journey of medical devices, we have gained valuable start-up business experience by overcoming challenges related to regulatory approvals, safety testing, documentation, funding, obtaining a reimbursement code, manufacturing to high-quality standards, and sales.

The Zamenix journey illustrates how academic research can translate into clinical impact and commercial success. The author hopes our experience will inspire robotic researchers to pursue the commercialization of their valuable research results.

Biography:

Dong-Soo Kwon is Emeritus Professor of the Department of Mechanical Engineering of KAIST and CEO of ROEN Surgical, Inc. He received the B.S. degree in 1980 from Seoul National University, the M.S.E. degree in 1982 from KAIST and the Ph.D. degree in 1991 from Georgia Institute of Technology, USA. He has been the vice President of Korean Innovative Medical Technology Society since 2019, Honorary chairman of the Korean Society of Medical Robotics, Honorary Chair of IEEE Ro-MAN 2023, and was President of Asian Society of Computer Aided Surgery (ASCAS), the program Chair of IEEE/RSJ international Conference on Intelligent Robots and Systems 2016, an academician at the National Academy of Engineering of Korea(NAEK), and IEEE Fellow.

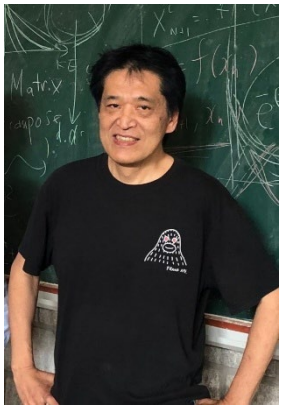
He received the 2023 Distinguished Service Award from the IEEE Robotics and Automation society, EAU2023 Best Video Abstract Awards First Prize, IROS Harashima Awards of Innovative Technologies in 2022, Top 10 Mechanical Engineering Technology of the Year 2019 from KFMES, Best Application Award and overall winner from Hamlyn Surgical Robot Challenge 2018, and the Harashima Mechatronics Award from ICROS 2009 and JTCF Novel Technology Paper Award from Amusement Culture from IROS in 2010. He also received double Korea Presidential Commendation (2018, 2022).

Invited Talk 1 (3F Meeting Room 32)

January 22 (Thursday), 09:00 - 10:45

OS25 ISBC: Individuality and Collectivity in Living Systems III

Organizers: Yasuhiro Hashimoto (University of Aizu, Japan)
Akiko Kashiwagi (Hirosaki University, Japan)



Community First Theory: How Collective Organization Generates Individual Diversity Across Biological, Robotic, and LLM Systems

Takashi Ikegami

The University of Tokyo, Japan

Across biological systems, individuality can emerge from the collective dynamics of communities. This is the central idea explored in this session, Community First Theory (CFT).

In Tetrahymena thermophila, single-cell RNA sequencing of 5,000 genetically identical clonal organisms revealed substantial transcriptional diversity within the population. Even after accounting for cell-cycle effects, gene expression clustered into six distinct groups enriched for ribosomal, mitochondrial, cytoskeletal, and peroxisomal proteins—indicating spontaneous specialization in metabolic and structural functions within a shared environment. These findings demonstrate that genetic identity does not guarantee phenotypic uniformity when organisms are embedded in collective contexts.

Similarly, in queen-less and male-less Pristomyrmex punctatus ant colonies lacking genetic variation, individual tracking uncovered two coexisting behavioral roles: ants forming dense clusters and others circulating between them. These groups alternated between deterministic exploratory movements and probabilistic clustering dynamics, with coordinated bursts resuming periodically. This behavioral heterogeneity was observed in honeybee hives as well. Together, these biological systems show that collective organization can amplify subtle individual differences, allowing distinct functional and behavioral traits to crystallize through interaction even in genetically homogeneous populations.

Recently we conducted a robot swarm experiment using ten identical phototactic robots under structured light conditions. Despite being physically and algorithmically identical, the robots exhibited spontaneous behavioral diversification over time. Subtle asymmetries in initial positioning and environmental symmetry breaking gave rise to distinct collective modes and coordination patterns, providing a minimal physical model for emergent individuality without pre-programmed heterogeneity.

A comparable process is observed in the digital ecologies of large-scale AI societies. We conducted multi-agent simulations using Large Language Model (LLM)-based agents that were initially undifferentiated—lacking predefined personalities, memories, or behavioral traits. These agents engaged in cooperative communication within group simulations, exchanging context-based messages in natural language. Through autonomous interactions, the agents spontaneously generated hallucinations and hashtags to sustain communication, which increased the diversity of words and topics within their exchanges. As conversations progressed, each agent's emotions shifted, and distinct personalities emerged and evolved as communities formed. Despite starting from identical initialization, the individuality of each agent arose not from its internal code alone but from distributed communication and mutual adaptation within the network.

Finally, using the framework of Partial Information Decomposition (PID), we characterize the informational advantages of developing individuality in terms of synergy and redundancy. This quantitative approach provides a unified description of differentiation processes across biological organisms, robotic collectives, and LLM-based agent societies, revealing a common information-theoretic foundation for emergent individuality.

Together, these observations support the Community First Theory, which posits that individuality arises from collective organization rather than existing prior to it. Differentiation is not a fixed design feature but a dynamic property of communities interacting with the same world—whether material, ecological, or informational. From Tetrahymena and ants to robots to networks of language models, individuality blooms from collectivity itself. The community gives birth to the very differences that sustain its function. This perspective offers new insights into the origins of diversity, autonomy, and intelligence in both natural and artificial systems.

References:

1. Hiroki Kojima, Akiko Kashiwagi, A & Takashi Ikegami (2024) Revealing gene expression heterogeneity in a clonal population of tetrahymena thermophila through single-cell rna sequencing. Biochemistry and biophysics reports, 38: 101720.
2. Norihiro Maruyama, Shigeto Dobata, Takashi Ikegami, Behavioral Analysis of Ant Colonies: Distinguishing Between Stochastic/Deterministic Modes and Global Behavior, AROBII-ISC-SWARM 2024.
3. Tomoyuki Atsushi Masumori, Norihiro Maruyama Takahide Yoshida Takashi Ikegami, From Swarm to Individual: Emergent Individuality in Light-Mediated Robot Collectives (submitted to Advanced Intelligent Systems 2025)
4. Ryosuke Takata, Atsushi Masumori and Takashi Ikegami: Spontaneous Emergence of Agent Individuality Through Social Interactions in Large Language Model-Based Communities. Entropy, 26(12), 1092 (2024).

Biography:

I have been working on the field of artificial life for more than 20 years. Evolution of genetic codes, mutation rates and cooperative relationships is one the main targets of my research. For example, complexity of coupled cognitive systems have been studied using dynamical recognizers and other recurrent neural (often embodied) systems. Recently, I am interested in constructing artificial life in the real world. To fruition the concepts developed through the study of artificial life, such as "autonomy", "enaction", "sustainability", and "evolvability", I have newly started several experimental and conceptual works, using an android (called Alter3), a large scale Boids model and other bio-chemical experiments.

Invited Talk 2 (3F Meeting Room 32)

January 23 (Friday), 09:00-10:45

OS26 ISBC: TONAL 2026

Organizers: Georgii Karelin (Okinawa Institute of Science and Technology, Japan)
Milan Rybar (Okinawa Institute of Science and Technology, Japan)
Moritz Kriegleder (Okinawa Institute of Science and Technology, Japan)
Luna Wang (Okinawa Institute of Science and Technology, Japan)

A novel formalization of the mind-body relation

Tom Froese

Okinawa Institute of Science and Technology, Japan

Cognitive sciences often appeal to intentions, reasons, motivations, and decisions to explain and predict behavior. However, such mental features are not directly observable in low-level observations, such as neural measurements.

From that, it follows that there is a conceptual tension between how agency is discussed in philosophy and psychology and what neuroscience can measure. If both aspects are important, how can they coexist? In other words, if something like agency matters and is not just an epiphenomenal "ghost in the machine", where is it in the signal? Perhaps it is not directly present as a variable at all, but only indirectly as a systematic pattern in deviations from expected activity. In the opposite case, if the eliminativistic account is correct, then cognitive science has no real object to study, except for some epiphenomenal statistical correlations and tendencies.

In this talk, a Participation Criterion will be introduced and explained: accounts of mental causation should specify what measurable difference the presence of mental efficacy would make compared with a counterfactual scenario in which mental efficacy is not present. The core hypothesis is that agency and consciousness may be observable, but only indirectly and in a model-dependent way, as context-aware departures from predicted dynamics, which will lead to an increased unpredictability in neural activity. In other words, it will be possible to predict neural dynamics from the observation of current and past moments; irruptions will be associated with moments of increased uncertainty of predictions (in a similar way to how modern weather forecasts also include uncertainty estimation). Irruption theory then treats these departures (designated as "irruptions") as structured, agency-linked variability.

Biography:

Dr. Tom Froese is a cognitive scientist with a background in artificial life and complex systems theory. His research addresses foundational questions regarding the origin and nature of the human mind by integrating philosophy of mind, computational modeling, and human subjects research.

https://www.oist.jp/research/research-units/ecsu/tom-froese

Invited Talk 3 (3F Meeting Room 32)

January 23 (Friday), 09:00-10:45

OS26 ISBC: TONAL 2026

Organizers: Georgii Karelin (Okinawa Institute of Science and Technology, Japan)
Milan Rybar (Okinawa Institute of Science and Technology, Japan)
Moritz Kriegleder (Okinawa Institute of Science and Technology, Japan)
Luna Wang (Okinawa Institute of Science and Technology, Japan)

When Decisions Matter: Entropy Dynamics in EEG as a Marker of Deliberate Choice

Milan Rybar

Okinawa Institute of Science and Technology, Japan

Decision-making is often studied by contrasting arbitrary choices (little consequence) with deliberate choices (reasoned and meaningful). EEG work has reported readiness potentials before arbitrary choices but not before deliberating (imporant and hard) ones, suggesting that the two kinds of decisions rely on different neural dynamics.

EEG is a direct readout of electrophysiological activity. However, it is also a noisy, low-spatial-resolution measurement, so what matters may sit in the dynamics rather than in a single amplitude peak. Here we re-analyze EEG from a donation-choice task, focusing not on mean amplitudes but on variability. We applied several entropy and complexity measures and tracked their changes after stimulus onset.
The main result is a difference in post-stimulus entropy dynamics between conditions: deliberate choices showed higher entropy/complexity than arbitrary choices.

These findings shift “neural noise” from just nuisance to potentially useful signals. So, the “noise” might matter too. From this perspective, entropy-based EEG measures could serve as proxy markers of agential involvement, highlighting the constructive role of variability in adaptive decision-making.

Biography:

Dr. Milan Rybar joined the Embodied Cognitive Science Unit at the Okinawa Institute of Science and Technology (OIST) as a Postdoctoral Researcher in July 2025. Previously, he was a Postdoctoral Researcher in the Hatsopoulos Lab at the University of Chicago (April 2023–June 2025), focusing on brain–computer interfaces, and an AI Research Scientist Intern at Meta (June–September 2022). He earned his PhD in Computer Science at the University of Essex (Brain-Computer Interfaces and Neural Engineering Laboratory), where he investigated BCI-based communication by differentiating semantic concepts from brain activity. Earlier, he completed MSc (Artificial Intelligence) and BSc (General Computer Science) degrees at Charles University in Prague. His interests include AI, computational neuroscience, artificial life, evolutionary computation, bio-inspired algorithms, and multi-agent/complex systems.

https://milanrybar.cz/

January 21 (Wednesday), 09:00-10:15

3F Meeting Room 32

OS24 ISBC: Complex Systems and Dynamics

Organizer: Toru Ohira (Nagoya University, Japan)

OS24-1 Nuclear energy clarified by extending hyper-gourd theory based on momentum equation and quasi-stability principle

Kaishin Tanaka and Ken Naitoh
(Waseda University, Japan)

In our previous reports [1-4], the hyper-gourd theory, is derived, which is based on momentum equation describing time-dependent deformation motions of biological particles such as molecules hydrated like nitrogenous bases and amino-acids and also cells and abiological particles like lepton, bosons, and atoms, while applying the quasi-stability principle proposed by us. The theory of us reveals both symmetric and asymmetric divisions of atoms, i.e. bimodal frequency distributions plotted against particle size. Emphasis is placed on the fact that the theory reveals symmetric division appears for trigger with high energy for atom, whereas asymmetric one occurs for lower trigger, although previous approaches based on energy conservation law such as Bohr [5] cannot do.

OS24-2 Extracting Plant Control Panel Operation Skills Using Eye Tracking and Work Videos

Hideo Miyachi[1] and Mami Kimura[2]
([1]Faculty of Information systems, Tokyo City University, Japan)
([2]Faculty of Economics, Nagasaki University, Japan)

Attempts to transfer the tacit knowledge possessed by skilled workers to novices using multimedia have been made for a long time. Advances in machine learning have made it possible to estimate human posture from moving images, allowing us to understand movements without attaching markers to workers. Eye-tracking systems have also become affordable. By combining external observation with eye-tracking information, it may be possible to infer a worker's intent and task content. This study recorded three types of video—monitor footage observed by workers in a recycling plant's control room, eye-tracking footage, and close-up footage of hand movements—and attempted to infer work intent through task analysis and interviews. Discussion using a display system allowing simultaneous viewing of all three video types revealed that workers control material flow on the conveyor by manipulating controls on the material supply side. This was identified as one of the key judgment skills required for this task.

OS24-3 Effects of Environmental Context Manipulation Using Multiple VR Environments on Memory

Hinata Onitsuka[1] and Hideo Miyachi[2]
([1]Graduate School of Environmental and Information Studies, Tokyo City University, Japan)
([2]Information Systems, Tokyo City University, Japan)

Human memory is known to be influenced by environmental context and spatial cues during learning and retrieval. However, it remains unclear whether environments presented in VR affect memory to the same extent as real-world environments. Therefore, this study aimed to examine the effect of differing numbers of rooms in a virtual reality (VR) space on word memory. Participants learned 100 words (3-character hiragana/katakana) via audio in a VR space under either a 1-room condition or a 4-room condition. Evaluation metrics included a free recall test and a questionnaire. Results showed no significant difference in the average number of correct responses among the 14 subjects. However, within the two groups—those who preferred the one-room condition and those who preferred the four-room condition—the average number of correct responses was significantly higher for the preferred condition in both groups.

OS24-4 Pursuit Curves Toward a Moving Object on a Known Path: Dynamical Analysis and New Formulation

Sota Yoshihara
(Graduate School of Mathematics, Nagoya University, Japan)

We investigate the one-on-one pursuit and evasion problem in terms of dynamical systems. J. C. Barton and C. J. Eliezer derived a set of simultaneous differential equations involving the positions of both players. These equations can be transformed into a dynamical system. The dynamical system involves two variables: the angular difference between the velocity vectors of the two players and the natural logarithm of the distance between them. It is possible to express the system by using a single complex variable whose magnitude and argument represent the relative motion. Numerical simulation results indicate two things. First, the difference in the pursuer's trajectory between a circular-path evader and an elliptical-path evader can be characterized by the existence of an equilibrium point in this system. Second, when the pursuer moves faster than the evader, catching up corresponds to the divergence of the exponent, at which point the simulation terminates. Previous studies directly use the distance between the two players as a variable. This leads to unphysical simulation behavior — for example, the pursuer may incorrectly begin to move away after contact.

OS24-5 Amplitude Enhancement Phenomena in Coupled Non-Autonomous Delay Differential Equations

Kenta Ohira[1], Toru Ohira[2] and Hideki Ohira[3]
([1]Graduate School of Informatics, Nagoya University, Japan)
([2]Graduate School of Mathematics Nagoya University, Japan)
([3]Graduate School of Informatics, Nagoya University, Japan)

This paper presents an exact analytical solution for a simple non-autonomous delay differential equation (DDE), representing one of the first explicit solutions known for this class of systems. Using a Fourier-transform-based approach, the solution is expressed as a superposition of Gaussian functions, allowing detailed analysis of transient dynamics beyond conventional numerical or stability-based methods. Building on this exact solution, the authors interpret the equation as a delayed self-feedback unit and investigate an extended system consisting of two such units. A key contribution of the study is the identification of a novel amplitude enhancement phenomenon that arises when the feedback structure is rewired from self-feedback to cross-feedback, while keeping all system parameters unchanged. Numerical simulations demonstrate that even when each individual unit produces only small signals, cross-feedback interactions can generate oscillatory wave packets with amplitudes enhanced by as much as eight orders of magnitude. Despite this dramatic amplification, the system remains asymptotically stable, with all trajectories eventually decaying to zero. The results highlight the critical role of delay and network structure in shaping transient dynamics and signal amplification. Possible applications include signal processing, information engineering, and physical implementations using delayed feedback circuits. The paper also outlines future directions, including deriving exact solutions for cross-feedback systems and experimental verification of the proposed enhancement mechanism.

January 21 (Wednesday), 09:00-10:15

3F Meeting Room 33

OS3 AROB: AI for Modeling and Optimization in Complex Systems

Organizers: Tomoyuki Hiroyasu (Doshisha University, Japan)
Hiroshi Furutani (Doshisha University, Japan)

OS3-1 Analysis of COVID-19 time-series data using two-step outbreak model

Hiroshi Furutani and Tomoyuki Hiroyasu
(Doshisha University, Japan)

This study analyzes COVID-19 time-series data from the sixth wave in Japan, which was driven by the highly transmissible but less severe Omicron variant beginning in late 2021. The analysis uses the Gompertz distribution, a growth and mortality model less commonly applied than the logistic distribution to describe the outbreak. The authors propose a two-step model of epidemic dynamics. In this model, the ratio of daily to cumulative cases, R(t)=u(t)/U(t), is transformed using L(t)=-ln R(t). Plotting L(t) over time reveals two distinct phases: (1) a small, slowly rising and falling mound, and (2) a linear region representing the main outbreak. The relationship between the mound's properties and the subsequent outbreak is examined. Data for all of Japan, as well as Tokyo, Osaka, and Okinawa, are used to show the results of Gompertz model analysis using the two-step model. This study demonstrates that the two-step model successfully indicates the starting date of COVID-19 outbreak.

OS3-2 An MM Algorithm for CNHD

Kensuke Tanioka
(Doshisha University, Japan)

When subject-specific functional connectivity network matrices constructed from fMRI data and binary labels are given, the Connectivity Network analysis method with discriminative Hub Detection (CNHD) has been proposed as a framework that simultaneously learns discriminative connectivity patterns and identifies hub nodes. In the original CNHD, parameter estimation is performed using the ADMM algorithm; however, introducing auxiliary variables and constraints increases the number of variables and complicates the estimation procedure. Therefore, in this paper, we propose an MM-algorithm-based estimation method and present a framework that enables comparable estimation without increasing the number of variables.

OS3-3 Simulations-Based Optimization of Compression Time in Sludge Dewatering using Filter Press Operating Parameters

Poltak Sandro Rumahorbo[1], Nobuhiro Yazawa[2], Hiroki Naoi[2], Satoshi Kondo[1], Yan Lyu[1],
Warut Timprae[1] and Shinya Watanabe[1]
([1]Muroran Institute of Technology, Japan)
([2]Tsukishima JFE Aqua Solution Co., Ltd., Japan)

Sludge dewatering is one of the most energy- and cost-intensive processes in water treatment plants. This study presents an analytical and simulation-based framework for optimizing the compression time of a filter press system in the sludge dewatering process, with the goal of minimizing operation time while achieving target of water content in the sludge known as cake moisture and its reduction rate. The focus is on the mechanical separation of solids and liquids under pressure, where cake moisture decreases nonlinearly as compression process progresses. Initially, rapid dewatering occurs due to high filtration rates, followed by a slower phase as the cake thickens and approaches its equilibrium moisture content. The physical behavior of the dewatering process following the exponential decay function and this model is derived to represent the relationship between compression time and cake moisture. This study proposed two simulations: using one day measurement data and only a few points of measurement data at the beginning stage. The estimation results show that the RMSE is relatively small, 0.59-2.39% and simulation results demonstrate that the proposed model can effectively estimate the optimum stopping time of compression for achieving specific target condition of cake moisture under varying pressure conditions.

OS3-4 Exploring Chemical Space with SELFIES-VAE: An Evaluation of GRU Effectiveness

Kyota Sakishima, Masashige Suzuki, Satoru Hiwa and Tomoyuki Hiroyasu
(Doshisha University, Japan)

The immense size of chemical space presents a major challenge in the discovery of novel drug-like molecules. Deep generative models, particularly variational autoencoders (VAEs), have emerged as a powerful approach for learning continuous molecular representations that enable efficient exploration and optimization. However, conventional SMILES-based VAEs often suffer from training instability and invalid molecule generation. In this study, we propose a GRU-based VAE framework using the SELFIES molecular representation, which guarantees syntactic validity of generated molecules. The model employs a multi-layer Gated Recurrent Unit (GRU) architecture in both the encoder and decoder to effectively capture sequential dependencies in molecular structures. Molecules are encoded as one-hot SELFIES sequences and mapped into a probabilistic latent space via the reparameterization trick. To explore this latent space, we apply the multi-objective evolutionary algorithm NSGA-II, optimizing drug-likeness (QED), synthetic accessibility (SA), and molecular diversity directly in the latent space. Experiments on a subset of the ZINC20 database demonstrate that the proposed approach achieves stable training, high reconstruction validity, and effective convergence to a Pareto front. While the generated molecules remain structurally simple, the framework provides a robust foundation for multi-objective molecular optimization and data-driven chemical space exploration.

OS3-5 Exploration of Chemical Space via Multi-Objective Genetic Algorithms with Molecule Fixing Crossover (MolFiX)

Sotaro Yamada[1], Satoru Hiwa[2] and Tomoyuki Hiroyasu[2]
([1]Graduate school of Life and Medical Sciences, Doshisha University, Japan)
([2]Department of Biomedical Sciences and Informatics, Doshisha University, Japan)

Generating novel, drug-like molecules remains a central challenge in computational drug discovery. In our previous work, wedevelopedtheMoleculeFixingCrossover(MolFiX),aSMILES-basedcrossoveroperatordesigned to preserve syntactic and chemical validity by addressing common errors such as unmatched ring closures, excessive valence, and broken branch syntax. In this study, we evaluate the effectiveness of MolFiX for multi-objective chemical space exploration, rather than its algorithmic design. We employ an NSGA-II–based multi-objective genetic algorithm (MOGA) using MolFiX as the crossover operator and consider three objectives: maximizing Quantitative Estimation of Drug-likeness (QED), minimizing Synthetic Accessibility (SA), and promoting molecular novelty. MolFiX is compared with a simple SMILES crossover without constraint handling and the MolFinder crossover. The results show that MolFiX consistently achieves a valid SMILES generation rate exceeding 90%, outperforming other methods. In addition, MolFiX produces broader and evenly distributed Pareto fronts, indicating improved exploration of trade-offs among drug-likeness, synthetic feasibility, and novelty. Analyses of generated molecules demonstrate favorable QED–SA trade-offs and structural diversity beyond local regions of chemical space. These findings confirm that MolFiX is an effective crossover operator for multi-objective molecular design, enabling robust and diverse chemical space exploration under realistic optimization settings.

January 21 (Wednesday), 09:00-10:30

B1 Meeting Room 1

GS1 Agent-based modelling

Chair: Kazushi Ikeda (Nara Institute of Science and Technology, Japan)

GS1-1 Research on the Enhancement of Precision of Battle Agent Model

Ryugo Okuyama, Masao Kubo and Hiroshi Sato
(National Defense Academy Yokosuka-shi, Kanagawa, Japan)

There is a growing movement to introduce Large Languadge Models(LLMs) into multi-agent simulations to overcome the challenge of homogenized interactions.In this study, we forcus on refinement of the 'Battle Agent Model' as one such example.Battle Agent is a multi-agent simulation that utilizes LLMs to recreate medieval battlefields; it has successfully reproduced unit-level behavioral changes resulting from eqipment differences between the English and French armies.On the other hand, its precision remains unexplored in several areas specifically the relationship between individual emotions and unit actions, the model's versatility, and its sensitivity to interactions with terrain.This paper reports the results of our efforts to address these challenges.

GS1-2 Evaluation of Lighting Arrangement Effects in Learning Spaces Considering Cognitive and Psychological Aspects of Lighting Color

Chhengnay Ich and Toshiyuki Asano
(Shonan Institute of Technology, Japan)

Designing effective learning spaces requires consideration of both cognitive and psychological factors. Lighting plays an important role in shaping perception, concentration, and comfort, yet evaluating its spatial effects often requires complex and costly visualization systems. This study proposes a simplified spatial simulation framework to examine how lighting design influences learning environments. Virtual reality (VR) experiments were conducted to compare blue (high color temperature) and orange (low color temperature) lighting conditions using electroencephalogram (EEG) measurements and Semantic Differential (SD) questionnaires. While EEG results show limited differences, subjective evaluations showed that blue lighting was associated with greater perceived concentration, whereas orange lighting was linked to increased comfort and relaxation. These perceptual findings were then incorporated into an agent-based simulation model implemented in NetLogo. The model explored how spatial zoning of work and recovery areas, as well as the inclusion of indoor plants, influenced long-term concentration patterns. Simulation results showed that optimized spatial configurations supported sustained attention and reduced cognitive strain. Overall, this study demonstrates the value of combining practical VR experiments with agent-based modeling to support evidence-based learning space design.

GS1-3 Simulation Speed-Up Method for Emergency Vehicle Route Selection —Evaluation in an Environment Simulating Real Roads—

Daiki Akutagawa[1] and Atsuo Ozaki[2]
([1]Graduate School of Information Science and Technology, Osaka Institute of Technology, Japan)
([2]Department of Information and Computer Science, Osaka Institute of Technology, Japan)

In recent years, while emergency vehicle dispatch requests have been on the rise due to an increase in infectious diseases and natural disasters, traffic congestion in urban areas has also become more severe. In this situation, the prolonged arrival times of emergency vehicles at the scene have become a social issue, making the establishment of technology to quickly calculate routes an urgent priority. This paper proposes a route selection method that uses multi-agent simulation technology to quickly generate reasonable routes related to emergency vehicles by creating and deleting dummies of emergency vehicle agents on each road. Furthermore, we compared the accuracy and processing time of the proposed method with Static and Dynamic Dijkstra algorithms and evaluated it in an environment simulating a real road network. The experimental results confirmed that the proposed method achieves higher selection accuracy than Static and Dynamic Dijkstra algorithms, validating the effectiveness of simultaneous multi-path exploration using dummy agents.

GS1-4 Traffic Flow Optimization and Congestion Mitigation Using Reinforcement Learning with a Traffic Simulator

Rikuto Ito[1] and Hiroyasu Inoue[1,2]
([1]Graduate school of Information Science, University of Hyogo, Japan)
([2]Center for Computational Science, RIKEN, Japan)

This study investigates the impact of state representation on the performance of reinforcement learning–based trafficsignal control. We focus on two widely used state representations: a lane-configuration-aware representation based on Occupancyand a lane-configuration-agnostic representation based on Wave. For evaluation, we apply the Multi-Agent Advantage Actor–Critic model. without modifying its learning framework. First, experiments are conducted on a regular grid network to examinethe fundamental characteristics of the two state representations under uniform conditions. Subsequently, the same framework isapplied to the Sannomiya area of Kobe City, which is characterized by heterogeneous intersections, mixed road types, and multi-lane configurations. In addition to evaluations under the training traffic demand, further experiments are performed under higherdemand conditions that are not observed during training. Through these experiments, we analyze the stability of traffic signalcontrol during peak demand periods and discuss how delay time characteristics relate to control performance in complex urbanenvironments.

GS1-5 Reinforcement Learning-Based Traffic Signal Control at Intersections Considering Vehicles and Pedestrians

Kuto Ishigami and Toshiyuki Asano
(Shonan Institute of Technology, Japan)

Traffic congestion remains a major issue in Japanese urban areas, and inefficient intersection signal control is a key contributing factor. To evaluate deep reinforcement learning (DRL) under more realistic conditions, this study incorporates pedestrian crossing behavior into a SUMO-based traffic signal control model. Focusing on a single intersection, we compare fixed-time control, MODERATO, and a DRL approach using Deep Q-Networks (DQN). DQN achieved the lowest average vehicle delay in 8 of 9 traffic scenarios and reduced delay by approximately 30–40% under low to moderate traffic volumes. However, under near-saturated conditions, turning-vehicle spillback caused MODERATO to outperform DQN in some cases, indicating limitations of DRL when traffic exceeds certain thresholds. To explore more scalable traffic management, the model was extended to a 3×3 multi-intersection network. Due to the enlarged state space and inter-intersection interactions, Proximal Policy Optimization (PPO) was adopted instead of DQN. Ongoing experiments compare decentralized and centralized control strategies, examining total delay, congestion propagation, and gridlock prevention. These results demonstrate that DRL remains effective even with pedestrian interactions while highlighting challenges under extremely heavy traffic.

GS1-6 Bringing the Dual-LLM Pattern to Practice in Google ADK for Deterministic AI Agent Security against Prompt Injection

Shuo Huang and Shin-Jie Lee
(Department of Computer Science and Information Engineering,
National Cheng Kung University, Taiwan)

Large Language Models (LLMs) are increasingly deployed as autonomous agents but remain highly vulnerable to prompt injection attacks that can induce unauthorized actions or data exfiltration. Probabilistic filtering cannot provide deterministic safety, motivating architectural defenses such as the Dual LLM pattern and the CaMeL framework. To make these designs deployable, we implement a Dual LLM–inspired security plugin within Google's Agent Development Kit (ADK). The plugin (1) isolates a privileged LLM (P-LLM) from a quarantined LLM (Q-LLM), (2) replaces untrusted content with UUID-based symbolic keys, (3) enforces structured, schema-validated exchanges, and (4) applies customizable pre-tool security policies via ADK callbacks. We recreated representative AgentDojo tasks in ADK and simulated 27 security test cases. The undefended baseline passed only 7 cases, with 20 compromised task completions. Dual LLM isolation improved robustness to 19 cases, blocking most injections. Remaining failures stemmed from malicious data extracted by delegated subtasks or reliance on untrusted data—issues mitigable through stricter policies or trusted data sources. These results show that Dual LLM architecture provides a practical and effective foundation for secure LLM-based agents.

January 21 (Wednesday), 09:00-10:30

B1 Meeting Room 2

GS6 Artificial life

Chair: Tatsuo Unemi (Soka University, Japan)

GS6-1 How Construction Material Properties and Costs Can Affect Coevolution of Morphology, Behavior, and Niche Construction in Soft Virtual Organisms

Miyu Ono, Takaya Arita and Reiji Suzuki
(Graduate school of Informatics, Nagoya University, Japan)

Organisms not only adapt to their environment but also modify it through niche construction, altering selection pressures on themselves and others. Using Evolution Gym, an evolutionary experiment platform for two-dimensional virtual soft robots, we investigated how physical properties (weight and stiffness) and costs of construction materials affect the coevolution of morphology, behavior, and niche construction strategies in soft-bodied virtual organisms. We adopted a valley-crossing task where organisms accomplish the goal by placing blocks on the field to modify the terrain as a form of niche construction. Valley-crossing experiments with four block types, under different placement cost settings, revealed that organisms evolve adaptive morphologies and behaviors tailored to material conditions, such as vertical actuators for traversing uneven terrain and horizontal actuators for pushing light blocks. Furthermore, we found that cost conditions promoting adaptive evolution depend on material properties; block weight serves as an implicit cost guiding efficient strategies in hard materials; and fundamental physical constraints in soft materials impose limitations that cost adjustments alone cannot overcome.

GS6-2 How Developmental Plasticity Costs Can Affect Plasticity-First Evolution in Soft Virtual Creatures

Ryunosuke Higashinaka, Takaya Arita and Reiji Suzuki
(Graduate School of Informatics, Nagoya University, Japan)

The plasticity-first hypothesis proposes that environmentally induced phenotypic changes precede genetic mutations, thereby driving adaptive evolution, yet the influence of plasticity costs remains unclear. This study investigates how the maintenance cost of developmental plasticity influences the plasticity-first evolution. To this end, we employed voxel-based soft robots with voxels capable of expressing different traits depending on environmental conditions, and evolved morphology and locomotion in changing environments by introducing maintenance costs proportional to the number of such plastic voxels. Our experiments revealed that the plasticity-first scenario was frequently observed, particularly in low-cost conditions, where neutrally evolved plasticity facilitated adaptation. Importantly, while high maintenance costs generally suppressed the number of plastic voxels, certain lineages still achieved adaptation through this scenario by utilizing minimal plasticity. Furthermore, phenotypic analysis demonstrated that maintenance costs act as a selection pressure that accelerates genetic assimilation, driving a rapid transition from plastic traits to genetically fixed phenotypes.

GS6-3 Analysis of Phenomena in Particle Lenia with Free Energy Principle

Taketo Shibata[1] and Kunikazu Kobayashi[2]
([1]Graduate School of Aichi Prefectural University, Japan)
([2]Aichi Prefectural University, Japan)

The Free Energy Principle (FEP) asserts that living organisms minimize variational free energy to adapt to their environments, regarding it as an upper bound on sensory surprise. This study investigates the potential of the FEP to promote open-ended evolutionary dynamics, characterized by sustained diversity and ongoing adaptation, within Particle Lenia, a particle-based artificial life environment. We implemented neural networks based on the Helmholtz machine as internal generative models within individual particles, enabling them to autonomously adjust their parameters through variational free energy minimization. Experimental results showed that the particles adapted to their environment and converged to a stable equilibrium. We further examined system dynamics under environmental fluctuations induced by sudden changes in population size. Notably, an increase in the number of particles induced pronounced entropy fluctuations, which ultimately led to enhanced parameter diversity. In contrast, population reduction led to a decrease in entropy. While previous research has emphasized the role of extinction events in accelerating evolution, our findings suggest that "growth events"—specifically population increases—play a more significant role in driving evolutionary dynamics within this system. These findings provide new insights into the conditions required to achieve open-ended evolution in artificial life systems.

GS6-4 Mind-as-BIN Hypothesis: A New Perspective on Artificial Mind

Toru Moriyama[1], Motoaki Iimori[2], Masao Migita[3], Hanna Saito[4], Taichi Haruna[5], Masanori Hiraoka[6], Yusuke Ide[7], Hiroaki Isobe[8], Shin-Ichiro M Nomura[9] and Kohei Sonoda[10]
([1]Shinshu University, Japan)
([2]Chuo University, Japan)
([3]Shiga University, Japan)
([4]Toyo University, Japan)
([5]Tokyo Woman's Christian University, Japan)
([6]Kochi University, Japan)
([7]Nihon University, Japan)
([8]Kyoto City University of Arts, Japan)
([9]Tohoku University, Japan)
([10]Ritsumeikan University, Japan)

Physicalism in analytic philosophy faces a persistent challenge: explaining the emergence of mind from matter without invoking strong emergence, which implies creation ex nihilo. Panpsychism offers an alternative by positing mental properties in fundamental particles, yet this approach encounters two major issues—the origin problem, concerning whether new prototypes must be posited as smaller units are discovered, and the combination problem, regarding how such prototypes aggregate into higher-order minds. Our research proposes a different framework grounded in comparative psychology: defining mind as behavioral unpredictability and introducing the Behavioral Inhibition Network (BIN) as its underlying mechanism. BIN suppresses competing impulses to maintain adaptive behavior but can autonomously release inhibition, generating unexpected actions such as wall-climbing in pill bugs under unfamiliar conditions. This unpredictability evokes agency and subjective experience, key attributes of mind. Extending beyond animals, the Mind-as-BIN Hypothesis posits BIN-like systems across all life and even non-living matter, offering a prototype-free model aligned with network-based neuroscience and recent expansions of cognition to plants and unicellular organisms. This framework provides a novel, interdisciplinary approach to the mind–matter problem.

GS6-5 Design and Morphology Analysis of Mobile Robots Based on Natural Selection

Daichi Yonemura, Kei Kondo and Masato Ishikawa
(The University of Osaka, Japan)

We propose a computational design framework for generating novel locomotion patterns using a non-Lamarckian, naturally selective algorithm. The design process separates morphological and controller design: body morphologies evolve over generations via mutation and selection, while controllers are treated as acquired traits learned individually for each robot. Morphology design includes structural topology determination and realization of specific link and joint configurations under minimal a priori constraints. Controller parameters are optimized via Bayesian optimization in a no-obstacle plane environment to maximize locomotion performance. 20 trials, each consisting of 20 generations, produced 60 representative locomotion patterns. Rolling strategies utilizing the entire body were most frequently observed, and specialized patterns such as sidewinding also emerged. Functional roles of individual links were interpreted a posteriori, but the resulting movements depended on controller-imposed periodicity, indicating that morphology alone did not fully determine locomotion. These findings demonstrate the potential for morphology-driven locomotion design under minimal constraints and highlight the interplay between morphology and control. Future work will focus on reinforcement learning frameworks that allow locomotion to emerge from morphology–environment interactions without periodicity assumptions, as well as quantitative analysis of evolutionary dynamics and functional differentiation across generations.

GS6-6 Plantbot: Integrating Plant and Robot through LLM Modular Agent Networks

Atsushi Masumori[1,2], Norihiro Maruyama[1], Itsuki Doi[1], John Smith[1], Hiroki Sato[1] and Takashi Ikegami[1,2]
([1]Alternative Machine Inc., Tokyo, Japan)
([2]The University of Tokyo, Japan)

We introduce Plantbot, a hybrid lifeform that connects a living plant with a mobile robot through a network of large language model (LLM) modules. Each module—responsible for sensing, vision, dialogue, or action—operates asynchronously and communicates via natural language, enabling seamless interaction across biological and artificial domains. This architecture leverages the capacity of LLMs to serve as hybrid interfaces, where natural language functions as a universal protocol, translating multimodal data (soil moisture, temperature, visual context) into linguistic messages that coordinate system behaviors. The integrated network transforms plant states into robotic actions, installing normativity essential for agency within the sensormotor loop. By combining biological and robotic elements through LLM-mediated communication, Plantbot behaves as an embodied, adaptive agent capable of responding autonomously to environmental conditions. This approach suggests possibilities for a new model of artificial life, where decentralized, LLM modules coordination enable novel interactions between biological and artificial systems.

January 21 (Wednesday), 09:00-10:15

B1 Meeting Room 3

GS12 Human-machine interaction and collaboration I

Chair: Sajid Nisar (Kyoto University of Advanced Science, Japan)

GS12-1 Development of robot design support system using multimodal LLM

Hayato Miura, Sho Yamauchi and Keiji Suzuki
(Future University Hakodate, Japan)

Recent advances have enabled robot design through dialogue with LLMs, yet their practical use remains limited due to output uncertainty and insufficient knowledge of hands-on fabrication skills. Much of the knowledge required for robot building is nonverbal and undocumented, making it difficult for LLMs to learn directly. This study addresses this issue by proposing a framework that uses a VLM to verbalize nonverbal fabrication information and makes it learnable for LLMs, enabling understanding of implicit technical knowledge. An experiment using VideoLLaMA3 examined how effectively nonverbal information could be extracted from a line-tracing car fabrication video. The process was divided into chapters, prompts were applied, and inference accuracy was evaluated. While visually clear elements such as subtitles and general components were recognized correctly, specialized parts like sensors and fine operations were often misinterpreted. To overcome these limitations, the study proposes integrating RAG with the VLM. By retrieving external knowledge relevant to the video, the system can enhance accuracy and compensate for information that visual analysis alone cannot capture. This approach is expected to improve the reliability and practicality of LLM-based support for robot fabrication.

GS12-2 Development of a LLM-based Interactive Agent in Word Wolf Game

Reon Ohashi, Joji Suzuki, Kazuya Tsubokura, Hibiki Sakurai and Kunikazu Kobayashi
(Aichi Prefectural University, Japan)

Recent advances in AI have shifted focus from perfect-information games to imperfect-information games involving communication. This study targets "Word Wolf," a social deduction game requiring players to identify a minority player through conversation while concealing their own status. We investigate whether Large Language Models (LLMs) can handle such contextually complex dialogue. First, we analyzed dialogue logs from human players, revealing that participants frequently use ambiguous backchannels to observe others without revealing sensitive information. Next, we developed Word Wolf agents using four LLMs: GPT-4o, Claude 3.7 Sonnet, llm-jp-3-13b-instruct3, and DeepSeek-R1-Distill-Qwen-14B-Japanese. We experimented with prompts incorporating insights from human gameplay and actual dialogue logs. Through 100 simulated sessions, results demonstrated that GPT-4o and Claude generated natural, human-like conversations, whereas the smaller models struggled. Notably, an ablation study indicated that providing raw human dialogue logs as references (few-shot prompting) resulted in more natural interactions than explicitly instructing the models with derived strategies. These findings suggest effective methods for applying LLMs to complex communication games.

GS12-3 Role-Playing Agent Construction via Automatic Prompt Optimization Using Embedding Representations

Yusuke Saito and Ryo Hatano
(Tokyo University of Science, Japan)

While fine-tuning is the dominant approach for constructing Role-Playing Agents (RPAs) that faithfully reproduce specific personas using Large Language Models (LLMs), it demands enormous computational costs and time. Conversely, prompt engineering offers a lightweight alternative but lacks systematic design guidelines. In this study, we address these challenges by proposing a framework that automatically optimizes prompts for building RPAs. Our method extracts prompts representing the persona's characteristics and past experiences. Using these prompts, an RPA generates responses. We then calculate the centroids of the distribution in the embedding space for both the target persona's utterance corpus and the set of responses generated by the RPA. We define the Euclidean distance between these two centroid vectors as the fitness function and explore the optimal prompt that maximizes this fitness using a genetic algorithm. This framework enables the construction of efficient and high-quality RPAs without requiring extensive computational resources and costs like fine-tuning.

GS12-4 Toward Ownership Understanding of Objects: Active Question Generation with Large Language Model and Probabilistic Generative Model

Saki Hashimoto[1], Shoichi Hasegawa[1], Tomochika Ishikawa[1], Akira Taniguchi[2], Yoshinobu Hagiwara[3,4], Lotfi El Hafi[4] and Tadahiro Taniguchi[4,5]
([1]Graduate School of Information Science and Engineering, Ritsumeikan University, Japan)
([2]College of Information Science and Engineering, Ritsumeikan University, Japan)
([3]Faculty of Science and Engineering, Soka University, Japan)
([4]Research Organization of Science and Technology, Ritsumeikan University, Japan)
([5]Graduate School of Informatics, Kyoto University, Japan)

Robots operating in domestic and office environments must understand object ownership to correctly execute instructions such as "Bring me my cup." However, ownership cannot be reliably inferred from visual features alone. To address this gap, we propose Active Ownership Learning (ActOwL), a framework that enables robots to actively generate and ask users ownership-related questions. ActOwL employs a probabilistic generative model to select questions that maximize information gain, thereby acquiring ownership knowledge to improve learning efficiency. Additionally, by leveraging commonsense knowledge from Large Language Models (LLMs), objects are pre-classified as either shared or owned, and only owned objects are targeted for questioning. Through experiments in a simulated home environment and a real-world laboratory setting, ActOwL achieved significantly higher ownership clustering accuracy with fewer questions than baseline methods. These findings demonstrate the effectiveness of combining active inference with LLM-guided commonsense reasoning, advancing robots' ability to acquire ownership knowledge for practical, socially appropriate task execution.

GS12-5 Comparing Performance of Automatic Speech Recognition Models for Conversational Robot Systems in Daily Environments with Individuals with Autism Spectrum Disorder

Ryota Fukui[1], Akishige Yuguchi[1], Yoshio Matsumoto[1], Hirokazu Kumazaki[2] and Yuya Okadome[1]
([1]Tokyo University of Science, Japan)
([2]Nagasaki University, Japan)

Recent advances in artificial intelligence, exemplified by large language models, have significantly enhanced the response generation capabilities in autonomous conversational robots. These systems are increasingly employed in everyday settings, and a promising application is the use of social robots to promote social skills growth for individuals with autism spectrum disorder (ASD). However, individuals with ASD show atypical prosodic patterns and other speech characteristics, making robust automatic speech recognition (ASR) essential for accurately interpreting their speech and avoiding inappropriate robot responses. Moreover, ASR for robots in real environments requires near-real-time latency. This paper evaluates different ASR models in terms of accuracy and latency, and investigates how recognition quality affects LLM-based robot responses, employing Japanese speech data collected in our preceding conversational experiments using an android robot. The dataset includes 45 sessions of 20 turns open-domain conversations with two participant groups: 13 undergraduate students and 10 individuals with ASD. We adopted six open-source, locally executable ASR models. Our results showed that character error rates were generally higher for ASD participants, indicating the difficulty of recognizing their speech. Comparative analyses of models suggested accuracy–latency trade-offs, informing ASR selection for real-world robot applications including ASD users.

January 21 (Wednesday), 09:00-10:15

B1 Meeting Room 4

GS18 Machine learning I

Chair: TBD

GS18-1 Offline Reinforcement Learning for Urban Emergency Driving: A Decision Transformer Framework for Mixed-Quality Datasets

Takuya Maniwa and Sachiyo Arai
(Chiba University, Japan)

Developing autonomous driving technology capable of safely managing sudden emergencies in urban environments is a critical challenge. Offline reinforcement learning (RL) has emerged as a promising solution, as it allows agents to learn effective policies exclusively from fixed, pre-collected datasets. However, offline RL faces two major hurdles: the inherent noise in datasets containing both expert and suboptimal trajectories, and a lack of "failure experiences." This absence of diverse data often leads to poor generalization and performance degradation in unseen, out-of-distribution (OOD) scenarios. To address these limitations, this study proposes a Decision Transformer (DT) framework trained on a mixed-quality dataset that intentionally integrates both expert and suboptimal trajectories. By leveraging a broader spectrum of experiences, the model gains the ability to distinguish between successful and unsuccessful outcomes, thereby enhancing its robustness in unpredictable environments. Simulation experiments demonstrate that our proposed method significantly outperforms standard DT models trained solely on expert data, as well as other offline RL algorithms. Our findings suggest that incorporating suboptimal data is essential for improving task success rates and ensuring reliable autonomous navigation within complex urban settings.

GS18-2 Multi-Vessel Route Planning under Limited Onboard Sensing via Asymmetric Reinforcement Learning

Tenyu Matsumoto, Fumito Uwano, Manabu Ohta and Donghui Lin
(Okayama University, Japan)

Recent advances in AI are driving autonomous navigation for small surface vessels, yet route planning and control that account for hydrodynamic disturbances remain challenging. Policies trained in idealized simulators often rely on environmental information that cannot be observed reliably in the real world because of hardware and sensing limitations, leading to severe degradation after deployment. We construct a virtual environment with Rankine vortices that generate realistic flow fields and propose a partially observable learning framework in which vortex information is supplied only with a fixed probability during training. The reinforcement-learning state includes vortex position, intensity, and rotation direction, but these variables are randomly masked so the agent learns behaviors that implicitly handle unobserved vortices without overfitting to explicit cues. We integrate the method into Implicit Quantile Networks (IQN) and two extended variants. Across navigation tasks, the resulting policies achieve higher target-reaching success and lower collision rates than baseline training schemes, indicating improved robustness to information asymmetry.

GS18-3 Learning Cooperative Autonomous Driving Using Local Common Rewards Based on V2V Communication Information

Tenta Suzuki[1], Johei Matsuoka[1], Tomohiro Harada[2], Kiyohiko Hattori[3]
([1]Tokyo University of Technology, Japan)
([2]Saitama University, Japan)
([3]Tokyo Denki University, Japan)

The realization of dedicated roads for fully automated vehicles is expected to reduce traffic congestion and accidents through cooperative group control utilizing V2V communication and sensors. However, manually designing control rules for such environments is prohibitively complex due to the vast number of vehicles and state transitions. Consequently, multi-agent reinforcement learning (MARL) has gained attention for enabling vehicles to autonomously acquire cooperative behaviors. While previous studies have shown that vehicles can develop emergent control through V2V communication without explicit rules, conventional individual rewards often suffer from the credit assignment problem. Specifically, since only colliding vehicles receive penalties, surrounding vehicles that contributed to the accident are not appropriately penalized, potentially encouraging risky driving. To address this, this study proposes a local shared reward mechanism that distributes penalties to both the colliding vehicles and their communicating neighbors. This approach encourages learning behaviors that

prioritize collective safety. Simulation experiments demonstrate that the proposed method reduces average lap times by approximately 15% and accident rates by 80% compared to individual reward structures.

GS18-4 MO-Decision Diffuser: Preference-Conditioned Diffusion for Multi-Objective Trajectory Generation in Automatic Train Operation

Cheng Liu and Sachiyo Arai
(Department of Electrical and Electronic Engineering,
Graduate School of Science and Engineering, Chiba University, Japan)

Modern automated train operation (ATO) must balance punctuality, energy efficiency, and ride comfort (stability) while satisfying strict physical and operational constraints. In real operations, expert driving demonstrations often concentrate into a few characteristic styles, leaving large portions of the feasible Pareto frontier under-explored. Weight-scanning reinforcement learning baselines typically retrain a separate policy for each preference vector and may produce solutions that are difficult to execute without highly detailed constraint modeling. We propose MO-Decision Diffuser (MODD), a preference-conditioned diffusion framework that learns once from expert-feasible trajectories and then generates velocity–time profiles for arbitrary trade-offs without per-weight retraining. MODD trains offline on normalized speed trajectories with contextual inputs such as speed envelopes and planned running time, and conditions generation on a normalized preference vector over the three objectives. To ensure executability, we enforce guardrails during sampling by clipping to empirical speed envelopes and limiting acceleration. Generated candidates are scored by a three-objective evaluator and filtered to form an empirical Pareto frontier. Experiments on real inter-station data show that MODD improves Pareto coverage and feasibility relative to weight-scanning PPO, filling gaps between clustered expert styles and providing a deployable, preference-switchable family of ATO trajectories.

GS18-5 Towards agricultural robots for autonomous navigation using foundation models and multimodal AI

Junsei Tanaka and Yoshihiro Sato
(Kyoto University of Advanced Science, Japan)

Farm autonomous navigation involves coupled, multi-layer challenges, and conventional pipelines tend to bloat as more exception-handling mechanisms are added. Although foundation models and multimodal AI (FM/MM) are promising, their adoption remains limited. This study asks whether farm autonomy will evolve as an extension of conventional approaches or shift toward FM/MM-centered architectures. From 67 papers, we extracted primary challenges and systematized them into five groups spanning perception/estimation and planning/control, showing that failures are amplified by cross-layer error propagation rather than a single bottleneck. These findings explain why conventional systems often require continuous, case-specific fixes across modules. Referring to long-tail discussions in autonomous driving, we organize two directions: hybridization that retains conventional modules to secure robustness and verifiability, and strengthening data collection, generation, and evaluation for rare cases. We further identify reasons for slow adoption—technical maturity, stepwise diffusion of advanced technologies in agriculture, and insufficient datasets and benchmarks—and finally present guidelines emphasizing shared data infrastructure, common evaluation protocols, and incentive design to sustain farmer participation.

January 21 (Wednesday), 13:00-14:45

3F Meeting Room 32

OS2 AROB: AI & Digital Transformation (DX) for Societal Systems - Explanation, Selection, Design & Evaluation

Organizers: Masanao Ochi (Oita University, Japan)
Masanori Shiro (National Institute of Advanced Industrial Science and Technology, Japan)
Shogo Matsuno (The University of Electro-Communications, Japan)

OS2-1 LLM Explanation Framework via Inter-Station Networks and External Factor Integration: Estimating the Origin of Boarding/Alighting Fluctuations and Explaining Their Propagation

Yuya Naruse and Masanao Ochi
(Oita University, Japan)

We present an explanation-ready pipeline integrating a learned inter-station network, a residual GCN, and an LLM agent to analyze railway ridership. Ridership is subject to complex periodicities and sudden shocks (e.g., weather, events) that propagate dynamically through the network. Previous works often treated stations in isolation, failing to explain the geographic origin and spread of these fluctuations. Our goal is to bridge structural forecasting with semantic interpretation. We construct a geometric k-NN graph and train a residual GCN incorporating cyclic time features. The model achieves a sharp improvement in macro-MSE over naive baselines. Graph diagnostics confirm high edge similarity, synchronous propagation, and robustness against hub ablation. Crucially, this learned network provides the necessary structural evidence for our LLM agent. By combining high-precision predictions with external factor grounding, the system generates accurate narratives describing not just why a fluctuation occurred, but where it originated and how it propagated across the railway network.

OS2-2 A Study on the Utilization of Failure Response Documents Using a Retrieval-Augmented Generation (RAG) Approach in City Gas Production Plants.

Kazuya Ono[1,2] and Masanao Ochi[1]
([1]Oita University, Japan)
([2]Oita Gas co.,ltd., Japan)

City gas production plants require rapid failure response to ensure safety and stable supply; however, current responses rely heavily on the experience of skilled operators. This study proposes an automated selection method for failure response documents using a Large Language Model (LLM) equipped with a Retrieval-Augmented Generation (RAG) mecha- nism, operating within a secure local environment. Specifically, we compared and verified whether ”information selection via retrieval” or ”information compression via summarization” is more suitable for fault diagnosis under context constraints. In our experiments, we evaluated five methods—the proposed method (with RAG), a non-RAG method (using summarization input), TF-IDF, BM25, and random selection—using 50 actual response documents and 10 queries, including unknown failures. The results showed that the proposed method achieved an F1 Score of 0.900, significantly outperforming the summarization method (0.400), which suffered from the loss of detailed information, as well as other methods. Furthermore, we demonstrated that even for unknown failures with no past response cases, high utility ratings were obtained from field operators if the proposed countermeasures were appropriate. This study demonstrates the superiority of RAG in fault diagnosis and highlights the practical utility of LLMs that cannot be measured solely by quantitative metrics.

OS2-3 Anonymity Evaluation Algorithm for Generating Explanatory Texts Targeting Individual Movement Histories

Ayumu Kai and Masanao Ochi
(Oita University, Japan)

Recent advances in IoT have enabled large-scale collection of sensor-based time-series data, but their complexity makes analytical results difficult for non-experts to interpret. This has increased demand for methods that translate complex outputs into natural language, with large language models (LLMs) emerging as a promising solution. However, while LLM-generated text improves interpretability, it also introduces new privacy concerns. Prior research has mainly focused on protecting either raw data, through techniques such as anonymization and Differential Privacy (DP), or the LLM inference process itself. In contrast, the privacy risk inherent in the generated natural-language output has not been directly evaluated. To address this gap, we propose a novel anonymity metric, N(D), inspired by the Akinator game, which quantifies the minimum number of questions required to uniquely identify an individual from a given description. Using mobility data, we generated textual summaries with an LLM and evaluated the effects of DP under different privacy settings. The results show that DP substantially increases N(D), indicating higher re-identification difficulty, and that N(D) sensitively captures variations in the DP parameter ϵ. Overall, this study bridges NLP and privacy-preserving data analysis by providing an intuitive framework for quantifying anonymity in LLM-generated text.

OS2-4 An Efficient Optimization Framework for Seismic Retrofit Design of Large-Scale RC Structures Using the Flower Pollination Algorithm

Shota Mine and Masanao Ochi
(Oita University, Japan)

Many public facilities, including wastewater treatment plants, have been in service for over half a century, and a significant number of them fail to satisfy current seismic design codes. Given that these facilities serve as critical infrastructure during disasters, the implementation of seismic retrofitting is an urgent priority. However, current retrofitting design relies heavily on the empirical judgment of skilled engineers. Furthermore, in densely populated cities, the structural scale of wastewater treatment plants tends to be large. Consequently, the number of possible retrofitting combinations becomes enormous, resulting in a drastic increase in the design burden. This study proposes an efficient optimization framework for the seismic retrofit design of large-scale reinforced concrete (RC) frame structures using the Flower Pollination Algorithm (FPA). By utilizing analysis results from the structural calculation software "SS7," the problem is formulated as an optimization task to minimize material costs subject to seismic performance constraints. This approach ensures both seismic safety and economic efficiency, thereby bridging the gap between theoretical optimization and practical design. In this study, wall placement is modeled as a discrete search space, and the FPA is employed to explore feasible retrofitting schemes that satisfy the required structural seismic indices. Numerical experiments on an existing RC structure confirmed that the proposed method efficiently identifies retrofit plans while ensuring both adequate seismic performance and economic efficiency. These results demonstrate the effectiveness of the FPA in addressing seismic retrofitting design problems.

OS2-5 A Topic-Aware Measure for Capturing Diverse Social Media Conflicts

Yusei Matae and Masanao Ochi
(Oita univercity, Japan)

Online conflicts shape beliefs, voting behavior, and information fairness on social networks, yet most polarization measures treat conflict as a single division and miss how tensions shift across overlapping issues. We present a topic-aware conflict metric that combines topic modeling with information-flow analysis to capture dynamic, cross-topic disputes. Using election-related posts on Bluesky, we build directed interaction networks from reposts, replies, and likes, estimate user topic engagement with BERTopic, and quantify disagreement via negative information flows within each topic. We then aggregate topic-specific signals, weighting them by users' levels of topic involvement, to obtain an overall conflict score. Comparisons with standard baselines, including SPIN and random-walk separation, indicate that the proposed method reveals fine-grained tensions—especially within narrow or cross-linked topics—that network-only approaches often overlook. Qualitative analyses further show that these detected cases correspond to substantively meaningful conflicts. By integrating topic awareness while keeping computation lightweight, our approach offers a practical and scalable tool for nuanced debate analysis and fairer assessment of information exposure during sensitive periods such as elections.

OS2-6 A Literature Filtering Method Using Central References for Evaluating the Disruption Index in Citation Networks

Ryosuke Motegi and Shogo Matsuno
(The University of Electro-Communications, Japan)

The Disruption Index (DI) is a widely used metric for quantifying the innovativeness of research. However, recent studies have pointed out that the DI is heavily influenced by "citation inflation," specifically the lengthening of reference lists over time. This inflation includes not only strategic citations but also "citation noise"—random or inaccurate citations that do not reflect the true flow of knowledge. In this study, we propose a literature filtering method to remove this noise and restore the accuracy of the DI. Specifically, we hypothesize that "central references," which serve as the intellectual foundation of a paper, form a dense structure within the citation network, whereas noise corresponds to peripheral documents. By applying centrality-based filtering to simulation data, we demonstrated that our method effectively reduces the bias caused by citation inflation and stabilizes the DI. This approach aligns with recent theoretical insights suggesting that the DI essentially measures the displacement of major prior research.

OS2-7 Test Data Generation Using Recurrence Plots

Masanori Shiro[1] and Masanao Ochi[2]
([1]National Institute of Advanced Industrial Science and Technology, Japan)
([2]Oita University, Japan)

This report introduced a method for data augmentation using surrogate data generated by recurrence plots to study social systems with AI, highlighting current challenges and potential solutions. Future work will focus on implementation and testing its application to some types of specific time series.

January 21 (Wednesday), 13:00-14:15

3F Meeting Room 33

OS10 AROB: Co-creation in research and education

Organizer: Kenneth J. Mackin (Tokyo University of Information Sciences, Japan)

OS10-1 Co-creation in inclusive design -developing audio-based action games-

Kenneth J. Mackin, Takeshi Fujiwara, Yoshihiro Kawano, Hideki Komagata and Keisuke Tanaka
(Tokyo University of Information Sciences, Japan)

Co-creation is the development of products where the user is included from the design stage. Tokyo University of Information Sciences (TUIS) recently established the Co-creation Laboratory, in which professors, students, and local government and firms work together to co-create a novel research and education environment. In the Co-creation Lab., students play a central role in designing the education program from which they learn. In this paper, we introduce initial results of one of the co-creation projects in progress at the Co-creation Lab. in TUIS. In this project, students designed and developed audio-based action games, which can be played by both blind and sighted students. The project members includes both blind and sighted students, and the goal of the project was to develop action games with inclusive design, in which both the blind and sighted students can compete fairly against each other. We give a description of the design decisions of the audio-based action game, explain the developed prototype program, and conclude with discussions on the co-creation project and future works.

OS10-2 A Consideration on Discrepancies Between Generative AI and Model Answers in Programming Education

Yorinori Kishimoto, Kenneth Mackin and Masanori Ohshiro
(Tokyo University of Information Sciences, Japan)

In programming education, it is often challenging to discern whether a student's submission is a direct output from a Generative AI (GenAI) tool. Furthermore, if a GenAI-generated solution during self-study deviates from the instructor's intended correct answer (model answer), the learning process may be compromised. Consequently, this study aims to analyze the correspondence between GenAI-generated solutions and the instructor's anticipated model answers for university programming assignments. By analyzing the differences between GenAI and model answers, it may be possible to assist in identifying whether a student's submission is GenAI-derived. Moreover, if a pattern in these discrepancies can be confirmed, it could highlight critical points for students to note during self-study and suggest avenues for support through effective prompt engineering.

OS10-3 A Distributed Cooperative Algorithm for the Practical Implementation of Drone Swarm Control

Xingjie Xie and Jonggeol Park
(Tokyo University of Information Sciences, Japan)

Multi-UAV systems have been widely applied to area coverage and cooperative missions due to their flexibility and efficiency. However, the failure of individual UAVs during mission execution can significantly degrade coverage performance and often requires manual intervention. To address this issue, this study proposes a distributed cooperative control framework with a dynamic task reallocation mechanism for multi-UAV systems. In the proposed approach, each UAV autonomously executes its assigned coverage task while periodically exchanging state information, including position and task progress, with neighboring agents via MAVLink communication. When a UAV failure is detected, the remaining UAVs cooperatively redistribute the unfinished tasks based on local information, without relying on a centralized controller or human supervision. The proposed method is implemented and evaluated in an ArduPilot Software-In-The-Loop simulation environment, where multiple UAVs perform area coverage missions under various failure scenarios. Simulation results demonstrate that the proposed framework can effectively maintain coverage continuity and improve system robustness compared with static task allocation strategies. These findings indicate that the proposed distributed control and task reallocation method is suitable for practical multi-UAV coverage applications in dynamic and uncertain environments.

OS10-4 Development and Practical Evaluation of a Learning Outcome Sharing Platform and Interview-Based Study to Promote Output-Oriented Learning among Children

Mikio Ito[1] and Yoshihiro Kawano[2]
([1]Tokyo University of Information Sciences Graduate School, Japan)
([2]Tokyo University of Information Sciences, Japan)

In the era of Society 5.0, self-directed and output-oriented learning has become essential for developing future-ready competencies. This study developed Manadasu, a cloud-based platform enabling children to share and accumulate learning products and examined its effectiveness through a community-based multigenerational programming classroom. The platform, built using Amazon Web Services, facilitates posting, sharing, viewing, and commenting on learning outputs such as videos and images. Five elementary school students participated in programming activities using Scratch. Semi-structured interviews and rubric-based assessments were conducted to evaluate learning outcomes across computational thinking dimensions including creativity, collaboration, and critical thinking. Results demonstrated significant growth in creativity and engagement. However, individual differences emerged in task understanding, particularly among younger participants, indicating developmental variations. Findings suggest that output-oriented learning enhances children's motivation and foundational abilities, while highlighting the necessity for instructional designs tailored to developmental stages. This research contributes insights for improving community learning activities and developing assessment frameworks for output-oriented education.

OS10-5 Vision-Based Online Landmark Mapping and Localization Using RGB-D Images

Guodong Zhu and Jonggeol Park
(Tokyo University of Information Sciences, Japan)

This paper presents a vision-based online landmark mapping and localization method using RGB-D images for robotic navigation, with a focus on UAV-oriented deployment scenarios. The system initializes a world coordinate frame at the first video frame and constructs an incremental landmark map from a set of high-response SIFT features. For each landmark, a 3D position is computed from pixel coordinates and depth measurements, and its appearance is represented by the associated descriptor. During online operation, features are matched between consecutive frames using FLANN-based KNN matching with a ratio test and a depth-consistency check to reject unreliable correspondences. Camera pose is then estimated from 3D–2D correspondences via a RANSAC-based PnP solver, enabling continuous trajectory estimation and map updates. Matched landmarks are updated with new observations, while unmatched valid features are promoted to new landmarks to expand map coverage. The proposed pipeline is implemented and evaluated with an Intel RealSense D455 RGB-D camera, demonstrating stable incremental mapping with real-time feedback and periodic map saving. The method provides a practical basis for future UAV integration in GPS-denied environments.

January 21 (Wednesday), 13:00-14:30

B1 Meeting Room 1

GS3 Artificial intelligence I

Chair: Noritaka Shigei (Kagoshima University, Japan)

GS3-1 Considerations on Reward Design in Reinforcement Learning for Multi-Agent Cooperative Tasks

Tatsuyuki Uto and Michiharu Maeda
(Fukuoka Institute of Technology, Japan)

Reinforcement learning (RL) includes the reward design of the crucial factors that influences the learning behavior of agents and the quality of its final policy. There are two types of reward design such as local rewards and global rewards. The agents learn quickly and may take the actions that differ from the overall goal in local rewards. Although the agents tend to encourage cooperative behavior in global rewards, the agents have a tendency to lead to local optima. This study introduces a hybrid reward design that combines local and global rewards with a weighted combination and adjusts the balance between them. We adopt the cooperative navigation in which three agents cover target landmarks while avoiding collisions. The local reward is defined as the negative distance to the nearest landmark, and the global reward is defined as the sum of negative distances between each landmark and its nearest agent. Both local and global reward designs impose an extra penalty when agents collide. We apply deep deterministic policy gradient (DDPG) and examine reward curves on different reward designs. Numerical experiments demonstrate that our model, which incorporates both local and global rewards and includes a constant term in both the initial and final training stages, outperforms models that utilize either local or global rewards alone.

GS3-2 Emergent Computational Structures for Reasoning in Language Models via a Simple Consistency

Jun-Hao Chen and Kazushi Ikeda
(Nara Institute of Science and Technology, Japan)

The development of complex, multi-step reasoning in artificial intelligence has predominantly relied on two resource-intensive paradigms: the explicit engineering of complex neuro-symbolic architectures and the massive scaling of model parameters. Both approaches present significant challenges in terms of computational cost, scalability, and generalizability, particularly for Small Language Models (SLMs), which often require expensive, data-hungry alignment techniques like process supervision to achieve robust reasoning. This paper introduces a third paradigm, inspired by the principles of self-organization and emergent computation observed in Artificial Life. We investigate whether a higher-order computational structure for logical reasoning can be induced to emerge within a standard SLM architecture, guided not by intricate design or massive scale, but by a simple, local, and task-agnostic learning constraint. Our methodology centers on a minimalist intervention during reinforcement learning. We employ a standard 3-billion-parameter model, Qwen2.5-3B, and optimize its policy using Group Relative Policy Optimization (GRPO), a computationally efficient, critic-free algorithm. The core of our approach lies in the reward function, which

imposes a programmatic consistency constraint. This constraint programmatically verifies that the model's final answer, enclosed in <answer> tags, is textually present as a substring within its generated chain-of-thought, enclosed in <think> tags. This simple rule creates a strong selective pressure, heavily incentivizing the model to generate reasoning pathways that logically and textually culminate in the final answer, thereby discouraging disconnected or "lazy" reasoning. This consistency reward is heavily weighted to prioritize the formation of internally coherent thought processes over mere outcome correctness. We validate our approach on challenging medical question-answering benchmarks, MedQA and PubMedQA. Our method achieves state-of-the-art performance for its model class, reaching 58.5% accuracy on MedQA and 76.2% on PubMedQA, representing absolute improvements of 10.7 and 8.9 percentage points over the supervised fine-tuned baseline, respectively. Notably, our 3B model outperforms larger models, such as the 4B MedGemma on PubMedQA, demonstrating that an efficient, emergent reasoning structure can be more effective than sheer parameter scale. Analysis of the learning dynamics provides further evidence of self-organization; the gradual stabilization of reward variance and completion length during training serves as a macroscopic signature that the model has converged upon a reliable and structured reasoning strategy. These results provide strong empirical evidence that a robust, functional reasoning capability can self-organize as an emergent solution to a simple learning objective. This work suggests a more scalable, efficient, and robust pathway toward advanced AI reasoning, shifting the focus from architectural design to the design of learning environments that foster emergent complexity.

GS3-3 A Comparative Study of Diffusion Models and Long Short-Term Memory Networks through the Reaction–Diffusion Equation

Vincent Lo, Kotaro James Nishimura and Kazushi Ikeda
(Nara Institute of Science and Technology, France)

In this work, we review the similarities between two machine learning algorithms, the Long Short-Term Memory (LSTM) networks and the Denoising Diffusion Probabilistic Models (DDPMs) through the Reaction-Diffusion (RD) equation, where morphogen concentrations vary depending on a diffusion with two different time constants and a reaction between them. On one hand, the LSTM networks have the interaction between long and short term memories, which is analogous to the RD mechanism with a discrete time framework. On the other hand, the DDPMs have both the drift and diffusion terms, which correspond to the RD mechanism. These facts mean that a general mathematical equation based on the RD mechanism can describe the two models. To empirically validate the theoretical mapping, a comparative experiment is designed to measure the internal dynamics of ``Reaction'' and ``Diffusion'' across the architectures. As results, their similarity and relationship were confirmed to some extent, although they are preliminary and need more studies in the future.

GS3-4 Kolmogorov-Arnold Message-Passing for Differentiable Causal Discovery

Brian Godwin Lim and Kazushi Ikeda
(Nara Institute of Science and Technology, Japan)

Graphical causality offers a principled framework for uncovering the data-generating process governing observational data. Nevertheless, identifying the causal graph structure, represented as a directed acyclic graph (DAG), often relies on domain knowledge and experts. To overcome this challenge, causal discovery methods aim to infer the causal DAG directly from data. Building on the additive noise model (ANM) paradigm, this paper introduces the Kolmogorov-Arnold message-passing (KAMP), an end-to-end differentiable framework for causal discovery. Specifically, KAMP integrates a differentiable DAG sampling model with a message-passing mechanism parametrized by Kolmogorov-Arnold Networks (KANs) to jointly learn both the causal graph structure and the associated functional causal relationships.

GS3-5 Quantitative Evaluation of Intrinsically Motivated Exploration Using a Light-Weight Task Enabling Fine-Grained Analysis

Momoka Yajima[1], Akane Tsuboya[2], Yu Kono[3] and Tatsuji Takahashi[1]
([1]Graduate School of Tokyo Denki University, Japan)
([2]Tokyo Denki University, Japan)

Random Network Distillation (RND) is an intrinsic motivation–based exploration method in reinforcement learning. But it has been reported to suffer from early collapse and insufficient variance of its intrinsic rewards. In this study, we design Scalable Pyramid, a task with a controllable structure of the state space and observation similarity. Using this task, we analyze the behavior of RND based Deep Q-Network while comparing it to a pure count-based intrinsic reward. We systematically vary the output dimensionality of the RND networks and the number of hidden layers in the predictor network, and compare the resulting behavior with that of the baseline count-based method. Our results show that increasing representational capacity keeps the intrinsic rewards at slightly higher levels; however, in all settings the intrinsic rewards still collapse rapidly within a small number of episodes, and the resulting improvement in final return is limited. These findings suggest that early collapse and variance deficiency arise not only from

properties of the environment but also from the design of RND itself, and highlights the need for multi-dimensional hyperparameter tuning and for exploring improved variants of RND.

GS3-6 Contribution-Based Modality-Switching Gate for Multimodal Action Recognition in Baseball Broadcast Video

Shin Okubo, Ryo Hatano and Hiroyuki Nishiyama
(Tokyo University of Science, Japan)

This work proposes a contribution-based modality switching framework for action recognition in baseball broadcast videos, in which commentary text is used only when it could improve prediction performance. We first construct a visual-only model (V-model) and a multimodal model that incorporates commentary text (VT-model), and define a contribution value Δ based on the loss difference between the two models. Δ is binarized to indicate whether the text is beneficial, and a Δ-Gate is trained to classify this contribution. In our experiments, the V- and VT-models achieved mean average precisions (mAPs) of 0.860 and 0.894, respectively, confirming that commentary text is effective for many cases. However, approximately 15% of samples exhibited $\Delta < 0$, indicating negative contribution in which commentary degraded performance. The proposed method uses the VT-model only for samples predicted by the Δ-Gate to have positive contribution, resulting in an mAP of 0.901, surpassing the VT-model. These findings suggest that our contribution-based dynamic modality selection is effective in handling the temporally asynchronous and semantically inconsistent linguistic information characteristic of sports broadcast videos.

January 21 (Wednesday), 13:00-14:45

B1 Meeting Room 2

GS8 Complexity

Chair: Hiroyuki Iizuka (Hokkaido University, Japan)

GS8-1 Do soldier crabs (Mictyris guinotae) recognize a robot as one of their own?

Ren Hiraguri and Yukio Pegio Gunji
(Waseda University, Japan)

Many animals form groups, known as swarms, in which individuals coordinate their movements with nearby conspecifics. Such collective behavior often involves approaching neighboring individuals and aligning movement directions. In this study, we focused on these characteristics and conducted experiments using Mictyris guinotae that forms swarms on tidal flats. These crabs are known to recognize other individuals visually and to approach simple moving objects that resemble crabs. In our experiment, a crab-shaped model robot was programmed to repeatedly move toward a specific crab individual, while the positions of the model and five real crabs were recorded using an overhead camera system. Based on the obtained coordinates, we analyzed the distance between the model and crabs, movement direction vectors of nearby individuals, and step length distributions. The results showed that crabs approached the moving model more frequently than a stationary model, and individuals near the model tended to partially align their movement directions with one another. These findings suggest that the crabs exhibit swarm-like behavior not only toward other crabs but also toward a moving artificial model.

GS8-2 Technological Fitness and Regional Economic Growth: Evidence from Japanese Prefectures

Rintaro Karashima[1] and Hiroyasu Inoue[1,2]
([1]Graduate School of Information Science, University of Hyogo, Japan)
([2]Center for Computational Science, RIKEN, Japan)

Technological knowledge has become a critical driver of regional economic performance in recent decades. This study investigates how the sophistication of technological knowledge relates to regional economic growth in Japan. Using 3.9 million corporate patent records from 1981 to 2015, we construct a bipartite network linking prefectures and technological fields and apply the Fitness–Complexity algorithm to measure the technological sophistication of each region. To evaluate its economic implications, we estimate fixed-effects panel models with Driscoll–Kraay standard errors, in which the dependent variable is the five-year real GRP per capita growth rate. The estimation results indicate that prefectures with higher Fitness values tend to achieve greater subsequent economic growth, reflecting the advantage of possessing more diverse and technologically advanced capabilities. This relationship remains robust after controlling for initial income levels, population density, patenting activity, unobserved regional

characteristics, and time-varying shocks. Lag and lead analyses further suggest limited endogeneity and do not support reverse causality. Overall, the findings highlight the importance of technological sophistication for long-run regional development in Japan.

GS8-3 Multilayer Propagation of Cross-Country Systemic Risk: A Hybrid Model with Horizontal DebtRank and Vertical Thresholds

Junhyun Chae[1] and Hiroyasu Inoue[1,2]
([1]Graduate School of Information Science, University of Hyogo, Japan)
([2]Center for Computational Science, RIKEN, Japan)

Financial shocks in cross-country systems are not confined to the financial sector but can propagate to the real economy and subsequently feed back into the financial system. In this context, the choice of analytical layer in country-level systemic risk assessment may substantially affect the interpretation of each country's systemic role. This study empirically examines how systemic risk assessments differ across layers by comparing a financial layer constructed from BIS Locational Banking Statistics (LBS) data with a real-sector layer constructed from UN Comtrade bilateral trade data, using a common exposure-based network framework. We develop a multilayer contagion model that integrates within-layer risk propagation based on the DebtRank algorithm with a threshold-based mechanism that activates cross-layer contagion once critical risk levels are exceeded. Using a balanced panel of 20 countries from 2000 to 2023, we compute layer-specific and multilayer-integrated systemic risk measures based on DebtRank propagation and CDS-implied default probabilities. The empirical results show that countries' systemic importance, risk transmission patterns, and relative rankings differ across layers and evolve over time. These findings demonstrate that country-level systemic risk assessments are structurally dependent on the chosen propagation layer and underscore the importance of a multilayer perspective when analyzing cross-country financial risk transmission.

GS8-4 Emergent Critical Dynamics in a Quantum-Entanglement-Inspired Cellular Automaton

Yoshihiko Ohzawa and Yukio-Pegio Gunji
(Waseda University, Japan)

Complex collective phenomena such as turbulence, morphogenesis, and social dynamics often exhibit universal critical features, yet the origin of this universality remains unclear. Cellular automata (CA) provide a minimal framework for studying such emergence, but classical CA typically require fine-tuned rules to generate criticality. In this study, we propose a one-dimensional CA whose update rules are defined by probability distributions inspired by quantum entanglement. The general problem addressed here is how quantum-like correlations embedded in probability distributions, rather than specific update rules, contribute to the emergence of critical dynamics. Here we show that introducing stochastic noise into this model actively promotes the emergence of 1/f-type fluctuations, serving as a clear signature of criticality. While both classical and quantum probability distributions exhibit noise-induced critical behavior, quantum distributions generate such dynamics more robustly in relatively high-noise regimes. These results indicate that noise is not merely disruptive but can be constructively exploited, and that the universality of critical phenomena may originate from the interaction between environmental noise and quantum-like probabilistic structures.

GS8-5 Semantic Reframing Through Hierarchical Operations: Toward a Cognitive Basis for Open-Endedness

Junki Kasano and Takashi Hashimoto
(Japan Advanced Institute of Science and Technology, Japan)

The realization and elucidation of open-endedness—the capacity for unlimited generation of novel diversity from finite interacting elements—is a central challenge in artificial life research. Although such open-endedness likely exists in human thinking, the cognitive mechanisms underlying it remain insufficiently understood. This study examined how linguistic hierarchical operations relate to diverse generation of others' mental contents in mind-reading to empirically explore a cognitive basis of semantic open-endedness. Forty Japanese native speakers completed two counterbalanced tasks: hierarchical operation task requiring recomposition of modifier relations in three-word compounds, and an inference generation task requiring diverse interpretations of others' mental contents from visual stimuli. A moderate positive correlation emerged between hierarchical operation efficiency and inference diversity, but only when hierarchical operation preceded inference generation, indicating an asymmetric effect. This relation was attenuated when controlling for measures related to divergent thinking. These findings identify semantic reframing grounded in linguistic structural recomposition as a candidate mechanism underlying cognitive open-endedness. The results provide behavioral evidence complementing simulation research on recursive combinations' role in generating structural diversity, extending it to semantic diversity and informing artificial life research when modeling artificial systems capable of generating open-ended semantic novelty.

GS8-6 Heterogeneity-Driven Regeneration of Collective Order in Soldier Crab Swarms

Ruijia Yang, Haruki Suzuki and Yukio Gunji
(Waseda University, Japan)

This study examines the macro-scale collective dynamics by introducing environmental heterogeneity, specifically alternating land and water zones, to crab swarms. The results revealed that the spatial absence of a uniform path forced the swarm to undergo a continuous cycle of splitting and reintegration. Notably, retrograde (reverse-moving) individuals emerged as critical triggers for this structural reorganization, generating a robust dynamic order that was absent in homogeneous control environments. These findings demonstrate that environmental heterogeneity functions as a constitutive mechanism for self- organization in biological collectives.

GS8-7 Toward Self-Organizing Error Taxonomies in Lightweight Language Models for Programming Education

Taku Matsumoto
(Hokkaido University of Science, Japan)

In programming education, logical errors that compilers cannot detect pose a significant barrier to learning. Traditional error classification systems depend on expert-defined categories, which require substantial manual effort and may not adapt to different problem sets. This study investigates whether lightweight language models can autonomously construct hierarchical error taxonomies through guided self-organization, without human-defined classification schemes. Using the 3-billion-parameter Qwen2.5-Coder-3B-Instruct as a lightweight language model, we propose a two-phase method. Phase 1 collects diverse error predictions for 44 programming problems through high-temperature sampling (T=0.7), obtaining 1,100 unique errors (8.3 times more than baseline). Phase 2 organizes these errors into hierarchical taxonomies via low-temperature generation (T=0.3), using a stepwise approach based on occurrence frequency to improve stability under limited model capacity. Experimental results show that diversity sampling achieves a diversity ratio of 84.16%. In Phase 2, taxonomy generation produced valid structured outputs for all high- and medium-frequency batches, and encountered only one parsing failure in the low-frequency per-problem stage, which was resolved by a single retry. The emergent taxonomy includes categories such as Runtime Errors (e.g., Division by Zero) and Logic Errors (e.g., Off-by-One Error), aligning with error patterns reported in prior studies. These findings suggest that guided self-organization can help lightweight models form meaningful error knowledge structures for educational support in local, offline environments.

January 21 (Wednesday), 13:00-14:45

B1 Meeting Room 3

GS19 Machine learning II

Chair: Kensuke Tanioka (Doshisha University, Japan)

GS19-1 Automatic Generation of Bow Movement for Violin-playing Robot ~ Results of Learning under New Formula for Rewards ~

Kenzo Horigome and Koji Shibuya
(Ryukoku University, Japan)

Recently, studies on human–robot communication have garnered significant attention. Musical communication is an important channel between humans and robots. We focused on robots playing musical instruments to build this channel. We focused on violins whose sounds are generated by bowing and fingering motions. In addition, we used reinforcement learning (RL) such that a violin-playing robot could automatically determine its bowing and fingering motions from different musical pieces. In this study, we extended the lengths of the musical pieces. In addition, we changed the formula for the rewards. We then gave a positive reward if the learning was successful. By contrast, we gave a negative reward if learning failed. The simulations were performed under these conditions. Consequently, the number of positive rewards obtained was greater than that of the negative ones, and the number of learning trials was directly proportional to the reward value of the obtained positive rewards. Thus, RL is apparently more feasible under different learning conditions in this study than in previous studies. This study is the first to achieve RL using a new formula.

GS19-2 Development of Motion Planning System for Violin-Playing Robot Using Reinforcement Learning —Sound Pressure Changes Based on Crescendos and Decrescendos—

Taiga Wakashiro and Koji Shibuya
(Ryukoku University, Japan)

The goal of this study was to use reinforcement learning to construct a system that automatically generates bow movements following composers' instructions for dynamic markings, such as the crescendos and decrescendos written in musical scores. Whereas conventional methods rely on explicit sound pressure target values in decibels, this study introduced a reward design based on relative changes in dynamics. In this system, target values such as "increase," "slight increase," "no change,""slight decrease," and "decrease" are automatically derived from MIDI data, enabling the system to interpret musical instructions directly from the score. The system uses the sound pressure equation modeled from bow speed and position to select bow speeds appropriate to the current bow position while adhering to hardware constraints such as joint angular velocity. Safety measures were incorporated to ensure the stable operation of an actual robotic platform. Using this framework, we generated bow movements for the first two measures of "Go Tell Aunt Rhody,"achieving some success in reproducing both crescendo and decrescendo patterns. These results show that by designing rewards around relative dynamic changes and directly linking score instructions to action generation, expressive bow movements can be achieved on hardware, realizing the composer's intended musical expression.

GS19-3 Natural-Language-Supported Reinforcement Learning via Large Language Model in Sparse-Reward Environments

Ryotaro Murakami[1], Fumito Uwano[2], Manabu Ohta[2] and Donghui Lin[2]
([1]Graduate School of Environmental, Life, Natural Science and Technology,
University of Okayama, Japan)
([2]Faculty of Environmental, Life, Natural Science and Technology, Okayama University, Japan)

Reinforcement learning (RL) is a framework in which an agent learns a policy through trial-and-error interaction with an environment, receiving scalar rewards as feedback. In sparse-reward environments, even optimal actions are often not immediately rewarded, which makes it difficult for the agent to distinguish good actions from random exploration. Large language models (LLMs) have strong reasoning ability and can potentially guide an agent's behavior, but their computational cost makes them impractical for real-time control. This paper proposes an RL framework that uses an LLM only as a temporary decision-support module during training. Concretely, the environmental information is embedded into a structured natural-language prompt, the LLM outputs an action guideline, and the RL policy is trained by imitating these LLM-generated behaviors. A key feature is an entropy-based scheduling strategy that adaptively determines the LLM-call probability from a normalized moving average of policy entropy. The LLM is called frequently when entropy is high and gradually phased out as the policy becomes more deterministic. Experiments on the MiniGrid DoorKey-16x16 task show that PPO alone almost never solves the environment, whereas the proposed LLM-assisted PPO reliably achieves near-optimal returns with LLM calls concentrated in the early to middle stages of training.

GS19-4 Improving Success Probability Estimation using a Weighted Loss Function and Frame Stacking in Safe Deep Reinforcement Learning

Shosuke Kawaguchi[1], Tohgoroh Matsui[2], Koichi Moriyama[1], Atsuko Mutoh[1], Kosuke Shima[1] and Nobuhiro Inuzuka[1]
([1]Graduate School of Engineering, Nagoya Institute of Technology, Nagoya, Japan)
([2]College of Life and Health Sciences Department of Clinical Engineering,
Chubu University, Aichi, Japan)

In recent years, it has become important for AI agents to autonomously avoid hazards in dangerous environments, such as disaster sites. Reinforcement Learning (RL) is widely used to learn autonomous behaviors, and the research field that focuses on learning risk-averse behaviors is known as Safe Reinforcement Learning (Safe RL). One existing method, SACEQ, aims to learn safe actions by biasing action values using ``success probability'' as a safety measure. However, SACEQ fails to accurately estimate this success probability. We identified that this failure stems from two main factors: the extreme data imbalance between success and failure samples, and the difficulty of capturing motion dynamics in continuous spaces. This paper proposes improving SACEQ by introducing a weighted loss function to address data imbalance and frame stacking to capture motion dynamics. Experimental results demonstrate that our full proposed method, which combines the weighted loss function and frame stacking, achieved a significant improvement, especially in Specificity. This result validates the effectiveness of our approach over the original SACEQ in successfully enabling the agent to correctly identify rare dangerous situations.

GS19-5 Self-Supervised Reward Densification for Multi-Objective Reinforcement Learning

Kota Minoshima and Sachiyo Arai
(Chiba University, Japan)

In multi-objective reinforcement learning (MORL), objectives can differ substantially in reward density: some provide frequent step-wise feedback, while others are sparse or weakly informative. In exploration-challenging settings, this density imbalance can bias learning toward dense objectives and hinder Pareto-front recovery. We propose a self-supervised reward densification framework for MORL that learns a vector-valued surrogate reward from online interaction without demonstrations. Our method stores trajectories, samples scalarization weights, and generates preference labels by comparing scalarized discounted vector returns between trajectory pairs. An ensemble vector reward model is trained with a Bradley--Terry loss so that predicted rewards yield preference-consistent trajectory scores, providing a denser learning signal for MORL policy optimization. We validate the approach on DeepSeaTreasure and on Highway-env as an application-oriented benchmark.

GS19-6 Stable Autonomous Control via Time- and Event-Triggered Hybrid Reinforcement Learning

Hayato Chujo and Sachiyo Arai
(Chiba University, Japan)

In recent years, research on optimization of control systems using online reinforcement learning, which simultaneously learns measures and controls with the measures, has been progressing. We focus on event-triggered reinforcement learning as an approach to optimize both control operations and time intervals. Compared with time-triggered reinforcement learning, which performs control operations at fixed time intervals, event-triggered reinforcement learning can solve the problems of instability caused by unnecessary control operations and increased control cost. However, the performance of event-triggered reinforcement learning tends to deteriorate in the early stages of learning due to the effect of initial settings, which is a cause of instability in control using online reinforcement learning.Therefore, we propose a combined time-triggered and event-triggered reinforcement learning method to improve the performance of event-triggered reinforcement learning in the early stages of learning. We also evaluate the performance of the proposed method by conducting computer experiments assuming the control of a heater.

GS19-7 Symbolic Music Classification Using MIDI2vec with Temporal Information

Zongyuan Shi[1], Atsuko Mutoh[1], Kosuke Shima[1], Koichi Moriyama[1], Tohgoroh Matsui[2] and Nobuhiro Inuzuka[1]
([1]Nagoya Institute of Technology, Japan)
([2]Chubu University, Japan)

Symbolic music offers direct access to the compositional structure of musical works, yet designing representations that capture both global symbolic relationships and local temporal patterns remains challenging. Existing approaches typically focus on either structural co-occurrence information or sequential dependencies, leaving a gap between graph-based representations and time-aware modeling. To address this issue, this study proposes a fused representation for symbolic music classification that integrates MIDI2vec, a graph-based structural embedding, with Sequence Graph Transform (SGT) features, which summarize temporal transitions using an exponential decay formulation. The two representations are combined at the feature level and evaluated using a multilayer perceptron classifier under stratified 10-fold cross-validation. Experiments demonstrate that temporal information can significantly improve performance over structural features alone when properly regularized by PCA, resulting in higher accuracy and macro-F1 scores. Additional analyses reveal that dimensionality reduction via PCA plays a crucial role in enhancing the discriminability of SGT features, with moderate compression yielding the best results. These findings highlight the importance of integrating structural and temporal perspectives in symbolic music representation. The proposed framework is computationally efficient, interpretable, and broadly applicable to downstream music information retrieval tasks.

January 21 (Wednesday), 13:00-14:15

B1 Meeting Room 4

GS25 Mobile robots I

Chair: Kazuyuki Ito (Hosei University, Japan)

GS25-1 Adaptive Policy Switching of Two-Wheeled Differential Robots for Traversing over Diverse Terrains

Haruki Izawa[1], Takeshi Takai[2], Shingo Kitano[2], Mikita Miyaguchi[2] and Hiroaki Kawashima[1]
([1]University of Hyogo, Japan)
([2]Takenaka Corporation, Japan)

Exploring lunar lava tubes requires robots to traverse without human intervention. Because pre-trained policies cannot fully cover all possible terrain conditions, our goal is to enable adaptive policy switching, where the robot selects an appropriate terrain-specialized model based on its current terrain features. This study investigates whether terrain types can be estimated effectively using posture-related observations collected during navigation. We fine-tuned a pre-trained policy using Proximal Policy Optimization (PPO), and then collected the robot's 3D orientation data as it moved across flat and rough terrain in a simulated lava-tube environment. Our analysis revealed that the standard deviation of the robot's pitch data shows a clear difference between these two terrain types. Using Gaussian mixture models (GMM), we evaluated terrain classification across various window sizes. An accuracy of more than 98% was achieved when using a 70-step window. The result suggests that short-term orientation data are sufficient for reliable terrain estimation, providing a foundation for adaptive policy switching.

GS25-2 Environment-Adaptive Dynamic Probability Exploration Strategy for Unknown Planets Using Tether-Connected Rovers

Clive Jancen Kawaoto[1], Johei Matsuoka[1] and Kiyohiko Hattori[2]
([1]Tokyo University of Technology, Japan)
([2]Tokyo Denki University, Japan)

The exploration of celestial bodies, such as Mars and the Moon, has intensified in recent years, necessitating methods for efficient survey of vast areas. Consequently, simultaneous exploration using multiple small rovers has emerged as promising. While our previous work introduced a tether-connected buddy rover system utilizing a fixed-probability strategy, this approach failed to adapt to varying environmental hazard levels, leading to frequent stuck incidents. To address this limitation, this study proposes a dynamic probability exploration strategy based on stuck frequency. The proposed method integrates a classifier-based basic strategy—where shallow grooves are traversed and deep ones are avoided—with a mechanism that dynamically adjusts the avoidance probability between 70% and 90% according to the frequency of being stuck. This allows for cautious navigation in hazardous zones and efficient data collection in safe environments. Experimental evaluations in three randomized environments compared the proposed method with conservative and fixed-probability strategies. The results demonstrate that the proposed method reduced travel distance by 16.7% and achieved a path efficiency of 48.9% compared to the conservative method, while also reducing stuck incidents by 20.9% compared to the fixed-probability approach. These findings suggest that the proposed strategy significantly contributes to the efficiency and safety of planetary exploration missions.

GS25-3 Sim-to-Real Acceleration Prediction Using Deep Learning for Shock Avoidance in Mobile Robots

Hiroto Shirono[1] and Kosuke Shigematsu[2]
([1]National Institute of Technology, Oita College, Electrical, Electronic Information Engineering, Japan)
([2]National Institute of Technology, Oita College, Department of Information Engineering, Japan)

Mobile robots traversing uneven terrain experience shocks and vibrations that can degrade stability and sensor reliability. This study proposes a deep learning framework to enhance safety by predicting the maximum resultant acceleration within a short future horizon. The model fuses multimodal time-series inputs including IMU data, velocity, command velocity, and depth images. A CNN extracts spatial terrain features from depth images, while an LSTM captures temporal dynamics to infer near-future physical loads.Experiments were conducted using datasets generated in Isaac Sim across slopes and step obstacles. The proposed model achieved real-time performance with an average processing time of 1.13 ms. For a 0.5 s prediction horizon, it attained an RMSE of 3.71 m/s^2. In contrast, a baseline that propagates the latest observed maximum acceleration produced an RMSE of 6.93 m/s^2. These

results demonstrate that the proposed method substantially improves prediction accuracy and can anticipate impact-related accelerations, supporting risk-adaptive motion planning for safer autonomous navigation.

GS25-4 A neural controller optimized via cuckoo search for an agricultural crawler vehicle

Yunosuke Tanaka[1], Kunihiko Nakazono[2], Takeshi Shikanai[2], Eiho Uezato[2] and Naoki Oshiro[2]
([1]Graduate school of University of the Ryukyus, Japan)
([2]University of the Ryukyus, Japan)

In this study, we propose a novel design method for an autonomous driving control system for crawler vehicles used in sugarcane harvesting. The core of the proposed approach is a neural controller (NC), which determines the vehicle's motion along a planned trajectory. To evaluate the performance of each NC, we define an evaluation function based on the sum of squared errors of position and speed at multiple evaluation points sampled at fixed time intervals along the trajectory. A smaller value of this evaluation function indicates better fitness and higher accuracy in following the desired path. To optimize the NC parameters, we employ the cuckoo search algorithm, a nature-inspired metaheuristic known for its efficiency in global optimization and robust search capability. Through extensive simulation experiments, we verify the effectiveness of the proposed method, demonstrating its potential to improve autonomous navigation, operational efficiency, and reliability in agricultural environments under varying conditions, while reducing manual intervention and overall operational costs.

GS25-5 Control System Designing a Neural Network Controller Optimized by Cuckoo Search for Crawler Vehicle

Ryoya Kikuchi[1], Kunihiko Nakazono[2], Takeshi Shikanai[2], Eiho Uezato[2] and Naoki Oshiro[2]
([1]Graduate school of University of the Ryukyus, Japan)
([2]University of the Ryukyus, Japan)

To address labor shortages and the aging population in the agricultural sector, particularly regarding sugarcane production in Okinawa, this paper proposes an autonomous driving control method for agricultural crawler vehicles. Navigating Okinawa's uneven and humid field conditions requires high-precision control capable of handling nonlinearity and environmental uncertainty. This study employs a three-layer Neural Network Controller (NC) optimized via Cuckoo Search (CS), a metaheuristic algorithm known for its superior search performance and convergence. A key innovation of this research is the implementation of a multi-stage evaluation function that separately assesses transient tracking performance (Interval A) and steady-state convergence (Interval B). To verify the stability and reproducibility of the optimization process, simulations were repeated 100 times with varying initial populations, comparing the best, median, and worst evolutionary trajectories. Furthermore, the system's generalization capabilities and robustness were evaluated by intentionally varying the vehicle's initial coordinates. The results demonstrate that the CS-optimized NC achieves stable path-following and effectively adapts to initial errors, providing a promising technological foundation for the advancement of smart agriculture in complex field environments.

January 21 (Wednesday), 16:05-17:05

3F Meeting Room 32

GS2 Artificial brain

Chair: Ken Saito (Nihon University, Japan)

GS2-1 The Synergistic Mind: Emergent Self-Regulation in the Humanoid ALTER3 with Concurrent Modular Agents

Takahide Yoshida[1], Norihiro Maruyama[1,2], Atsushi Masumori[1,2] and Takashi Ikegami[1,2]
([1]The University of Tokyo, Japan)
([2]Alternative Machine.inc, Tokyo, Japan)

We present a concurrent modular agent (CMA) architecture functioning as an artificial brain that instantiates Minsky's Society of Mind on a humanoid robot (ALTER3) deployed for months in a public venue. In this article, we report observations and analyses based on data obtained during several months of continuous public deployment at the Venice Biennale of Architecture. The CMA architecture comprises multiple LLM-based modules running concurrently and exchanging messages through a shared vector store. This complex and dynamic architecture facilitates engaging, non-repetitive interactions between visitors and Alter3, emerging in life-like behavioral patterns. Behavioral analyses show that when all modules are active, ALTER3 exhibits enhanced autonomy, emotional valence, and physical expressivity compared with a dialogue-only ablation

GS2-2 Embodied Cross-Modal Imagination: Learning Multisensory Cognitive Maps from Sound, Touch, and Vision

Itsuki Doi[1,2] and Takashi Ikegami[1,2]
([1]The University of Tokyo, Japan)
([2]Alternative Machine Inc., Japan)

Biological perception is inherently multimodal: agents interpret and anticipate sensory events by integrating vision, audition, touch, and self-motion signals. While predictive learning has been extensively studied in unimodal settings—particularly vision—real-world perception, imagination, and recall often rely on synergistic interactions across modalities, where combinations can give rise to representations not evident in any single modality alone. In this study, we train an embodied neural network to predict next-step sensory observations from synchronized multimodal inputs. We analyze how multimodal predictive learning shapes unified latent representations, focusing on whether specific combinations of modalities (e.g., vision and audio together) produce structured spatial codes or emergent features that cannot be learned from any single modality alone. We further examine cross-modal prediction—whether the network can generate plausible sensory expectations in one modality from information available only in others, revealing the generative capacity of unified representations.

GS2-3 Reusing Low-Level Spatial Representations for Higher-Level Representations in a Shared-Module RNN

Hiroyuki Iizuka and Wataru Noguchi
(Hokkaido University, Japan)

We propose a neural network model that acquires hierarchical spatial cognition through predictive learning from visuomotor experience. The model consists of hierarchically connected recurrent neural network (RNN) modules that share parameters across layers, enabling the self-organization of multi-scale spatial representations without explicit hierarchical supervision. A simulated mobile robot learns to predict future visual inputs while navigating a multi-room environment. Analysis of internal states reveals that lower-level modules represent local spatial positions within rooms, whereas higher-level modules encode the global arrangement of rooms, with structurally similar spatial organizations across scales. These results demonstrate that hierarchical spatial representations can emerge from shared neural mechanisms operating across scales through predictive visuomotor learning.

GS2-4 SUGAR: A Physically Grounded Semantic Unit Generating Affective Responses

Christian Francesco Carlino
(Independent Researcher, Italy)

The origin of meaning, life, and consciousness remains a fundamental open problem. Existing explanations, ranging from chance-based emergence to design-based models, do not clarify how information, coherence, and proto-semantic structure could arise within natural physical substrates. To investigate this gap, the Hydro-Magnetic Catalyst (HMC) was developed as a minimal physico-chemical system in which a gelled saline matrix is exposed to weak oscillating magnetic fields (1–60 Hz). The HMC functions as a symbolic demodulator: continuous voltage fluctuations generated by field–water interaction are converted into discrete transitions (A, C, G, T), formally reminiscent of biological coding strategies. Building on this mechanism, SUGAR (Semantic Unit Generating Affective Responses) extends HMC output through Shannon entropy, Lempel-Ziv complexity, and the Quantillium (QI) index, enabling discrimination between structured and unstructured stimuli, the emergence of simple affective states, and a compact form of semantic memory. SUGAR introduces a distinct perspective by emphasizing two physical substrates common to all living systems, structured water and low-frequency electromagnetic oscillations, rather than neuron-centered explanations. Despite its simplicity, the system shows that organized symbolic activity and meaning-bearing states can arise from field–matter dynamics alone, offering a minimal and reproducible framework for studying proto-semantic behavior and artificial consciousness.

January 21 (Wednesday), 16:05-17:20

3F Meeting Room 33

GS4 Artificial intelligence II

Chair: Brian Godwin Lim (Nara Institute of Science and Technology, Japan)

GS4-1 High-Fidelity Respiratory Sound Augmentation for Robust Classification: Leveraging StyleGAN2 and Autoencoder Pre-training

Takehiro Hirasawa[1], Yasumasa Tamura[2], Kaoruko Shimizu[2], Satoshi Konno[2] and Masahito Yamamoto[1]
([1]Faculty of Information Science and Technology, Hokkaido University, Japan)
([2]Faculty of Medicine, Hokkaido University, Japan)

Developing wearable-based automated auscultation systems is vital for remote patient monitoring, especially for respiratory diseases. However, deep learning application is hindered by the scarcity of high-quality, annotated abnormal respiratory sounds, which often leads to poor model generalization. While StyleGAN2 can synthesize realistic data, utilizing it for supervised classification introduces risks regarding label validity. To address this, we propose a feature learning framework. First, we modified StyleGAN2 to support rectangular inputs (128x512), ensuring the preservation of high-temporal-resolution characteristics critical for respiratory sounds, and validated it using a domain-specific protocol with PANNs and MMD. Subsequently, using 320,000 synthetic Mel-spectrograms, we pre-trained an Autoencoder to learn robust acoustic priors without label dependency. This encoder served as a fixed feature extractor for binary classification (Normal vs. Crackle) on a limited clinical dataset (N=37) collected via a wearable device. Our method achieved a Macro F1-score of 0.9499 and AUC of 0.9369, significantly outperforming a scratch model (F1: 0.3075). This demonstrates that transferring knowledge from massive synthetic data via unsupervised learning effectively overcomes data scarcity, paving the way for reliable automated diagnosis in home-care settings.

GS4-2 Improving CETNet-Based Temporal Action Segmentation through Downsampling-Based Data Augmentation in Laparoscopic Surgery Videos

Haruka Eto[1], Tsubasa Hidaka[1], Kanta Kubo[1], Kan Tanabe[2], Rara Deguchi[1], Kenji Baba[2], Naoki Kuroshima[2], Masumi Wada[2], Mashiho Mukaida[1], Takao Ohtsuka [2], Satoshi Ono[1] and Noritaka Shigei[1]
([1]Graduate School of Science and Engineering, Kagoshima University, Japan)
([2]Graduate School of Medical and Dental Sciences, Kagoshima University, Japan)

This paper proposes a Multi-Offset Downsampling (MODS) strategy for Temporal Action Segmentation (TAS) in laparoscopic surgery using the Cross Enhancement Transformer for Action Segmentation (CETNet). In surgical skill assessment, TAS plays a key role in recognizing fine-grained actions from videos; however, CETNet requires substantial memory, making downsampling (DS) unavoidable for long surgical sequences. Standard fixed-offset DS reduces temporal resolution and partially removes frame-level information, which can degrade the recognition of short but clinically important actions. To address this issue, MODS introduces multiple offset patterns during training to increase temporal coverage and data diversity without modifying the inference pipeline. We evaluate the method across different DS rates, focusing on both overall recognition accuracy and class-wise detection performance. Experimental results demonstrate that MODS improves accuracy by enhancing generalization and effectively reinforces the detection of surgically critical action classes such as “Hemostasis”, as well as relatively short-duration actions like “Cut”, “Clip”, and “Sharp Dissection”, which are often overlooked by conventional fixed sampling.

GS4-3 A Robust Multi-level Vision-only Framework for Few-shot Anomaly Detection

Yusuke Okido, Ryo Hatano and Hiroyuki Nishiyama
(Tokyo University of Science, Japan)

Few-shot anomaly detection constructs a normality model from a limited number of normal samples and detects deviations as anomalies, playing an important role in industrial visual inspection. Although patch-based methods using high-quality features from self-supervised learning achieve strong performance, detection of fine-grained anomalies can become unstable due to variations in illumination and luminance distributions, as well as distance evaluation designs relying on single-layer features. In this study, we propose a training-free, patch-based anomaly detection method using DINOv2 [1] features and investigate design-level improvements to stabilize fine-grained anomaly detection. Specifically, we introduce luminance correction, rotation-based augmentation to compensate for insufficient coverage of the normal distribution under few-shot conditions, local smoothing of patch-level anomaly scores, and multi-scale feature fusion to accommodate diverse anomaly scales. The proposed method requires no additional training and maintains a simple method, while achieving AUROC of 97.1%, F1-score of 96.5%, and AP of 98.6% under the 1-shot setting on the MVTec-AD [2] dataset, demonstrating consistent improvements over existing DINOv2-based methods. These results indicate that, in few-shot anomaly detection, preprocessing and the design of feature utilization and scoring are crucial for stabilizing fine-grained anomaly detection, in addition to the quality of feature representations themselves.

GS4-4 Critique-Guided Haiku Revision and Composition with Large Language Models

Soichiro Yokoyama, Tomohisa Yamashita and Hidenori Kawamura
(Hokkaido University, Japan)

Haiku composition requires not only adherence to formal constraints but also the effective conveyance of both scene depiction and emotional expression. While large language models (LLMs) have shown promise in automatic haiku generation, most existing approaches treat evaluation as a post-hoc scoring or selection process, limiting their ability to reflect human creative practice. This study proposes a critique-guided haiku revision framework that integrates interpretation, evaluation, and generation into a single creative loop. Given an input haiku, the system interprets key lexical elements and generates explicit textual critiques describing the depicted scene and emotional tone. These critiques guide iterative revision, producing rewritten candidates that preserve core imagery while improving clarity and coherence. All processes are conducted in Japanese using prompt-based interactions with a general-purpose LLM, combined with rule-based filtering for formal validity. Expert evaluations were conducted from three perspectives: individual haiku quality, relative quality assessment in a haiku-workshop style setting, and direct paired comparison before and after revision. Results show that critique-guided revision improves grammatical quality and emotional expression, while revealing a trade-off with scene depiction. Moreover, although overall quality improves through revision, a substantial gap remains between revised haiku and human-authored award-winning works, highlighting current limitations of LLM-based revision.

GS4-5 Finetuning an LLM on Structured Snapshots for Grounding Web UI Elements via Spatially-Enriched Training Data

Ai-Ni Li and Shin-Jie Lee
(National Cheng Kung University, Taiwan)

Traditional methods for locating web elements, such as XPath and CSS Selectors, are fragile under page revisions and costly to maintain. Semantic-oriented locators (e.g., Playwright's getByRole API) improve readability but remain limited by predefined rules and a shallow understanding of natural language. With the rise of AI-driven Web Agents, grounding UI elements from natural language has become increasingly critical. Playwright MCP now provides developers with two distinct grounding paradigms—visual-based (screenshot-driven) and structured information–based (DOM-driven)—allowing them to choose the method best suited for their application. Different types of Web apps may benefit from different approaches. Our work does not seek to prove that structured text-based grounding is superior to visual grounding. Instead, we focus on enhancing the structured, text-oriented grounding path, so that when developers opt for this method, they achieve significantly stronger results than with current baselines. To this end, we introduce a fine-tuned LLM trained on spatially enriched structured data. In this setting, element descriptions are augmented with spatial descriptors derived from bounding boxes and relative positions, enabling the model to jointly learn semantic and positional cues. This approach strengthens the capability of text-based grounding without requiring new vision-language models, making it practical to integrate into existing workflows built on structured snapshots. To validate this approach, we designed a chained workflow: a general-purpose LLM first attempts grounding on structured snapshots, and unresolved cases are routed to our fine-tuned model. Experiments on 10,000 test samples show that while a baseline LLM (LLaVA without visual inputs) achieves only 14% accuracy, our fine-tuned model recovers 88% of previously failed cases. These results highlight the effectiveness of spatial enrichment in structured, text-based grounding, underscoring its potential as an independent and scalable direction for integrating LLMs into Web Agent systems.

January 21 (Wednesday), 16:05-17:35

B1 Meeting Room 1

GS9 Control techniques

Chair: Fusaomi Nagata (Sanyo-Onoda City University, Japan)

GS9-1 Comparative Study on Location Estimation in Mini 4WD AI

Kotaro Tanaka, Tomoharu Nakashima and Yoshifumi Kusunoki
(Osaka Metropolitan University, Japan)

In this paper, we propose a method of location estimation using a geomagnetic sensor for the automatic control of a Mini-4WD AI car. A circular model is constructed from on the horizontal geomagnetic data acquired during driving, and the car orientation is estimated using the angle data. Then, the car orientation one step ahead is predicted using reservoir computing for the angle data, and the position is estimated based on the predicted orientation.

GS9-2 High-precision speed control of a hydraulic cylinder based on piston speed and external force feedback

Sota Umemoto[1], Koichi Osuka[2], Sugie Toshiharu[1], Akira Ishii[3], Yasuo Wakabayashi[3] and Masato Ishikawa[1]
([1]The University of Osaka, Japan)
([2]Osaka Institute of Technology, Japan)
([3]Komatsu Ltd., Japan)

In this study, a control system with feedback of external force and piston speed is proposed to achieve high-precision speed tracking under large and fluctuating external forces. A mathematical model of the hydraulic system and a simulator were developed to analyze how external forces influence piston speed. Based on this analysis, an external force feedback controller was designed, which was combined with speed feedback. In the external force feedback loop, the inverse dynamics derived from the identified transfer function was used to compute the valve input command based on the target speed and the estimated external force. This loop primarily compensates for large-scale external disturbances. In the piston speed feedback loop, sliding mode control was applied so that the system trajectories converge along a predefined sliding surface toward the target speed. This control law compensates for vibrations caused by disturbances and modeling errors, thereby improving the robustness of the overall control

system. The proposed control law was validated through both simulations and experiments, demonstrating accurate tracking performance with respect to the desired speed in all test conditions.

GS9-3 Obstacle Avoidance Control Based on Minimum Snap Trajectory for Autonomous Vehicles

Ryoma Hosoda and Ivan Tanev
(Graduate School of Science and Engineering Information and Computer Science,
Doshisha University, Kyoto, Japan)

As autonomous driving advances, control strategies that ensure both safety and ride comfort have become increasingly important. Although lateral acceleration and jerk are commonly used as comfort indices, lateral snap has received limited attention despite its strong influence on vehicle stability and passenger comfort, especially during emergency maneuvers. This study proposes a minimum-snap servo-control model that regulates steering commands using a smooth time-varying gain embedded in a proportional–derivative framework. The controller parameters are optimized using a Genetic Algorithm (GA), where the fitness function evaluates both the aggregate lateral snap and the precision of maneuver. To enable real-time application under unforeseen driving conditions, a machine-learning-based predictor is developed using GA-optimized solutions as training data. The proposed method is validated in a car simulator on a two-lane straight road with a stationary obstacle, under various speeds and time-to-collision conditions. Compared with a clothoid-based baseline, the proposed controller reduces lateral snap by approximately a factor of about 13 while maintaining smooth trajectories and stable steering behavior. These results demonstrate the effectiveness of the proposed snap-minimization approach for emergency obstacle avoidance.

GS9-4 Proposal of deformable pump-shell for compact and anti-fouling hemodialysis system

Mamiko Matsuo[1], Takehito Kikuchi[2], Yoshihiro Tange[3] and Tadashi Tomo[3]
([1]Graduate School of Engineering, Oita University, Japan)
([2]Faculty of Science and Technology, Oita University, Japan)
([3]Faculty of Medicine Oita University, Japan)

Wearable hemodialysis unit is one of the state-of-the-arts technologies for continuous 24/7 kidney replacement therapy and has great potential to enhance the quality of life for patients with end-stage kidney failure. However, the fouling of the dialysis filter is a significant consideration for such wearable systems owing to its restricted flow speed. The objective of this study is to develop a compact roller pump system that can prevent filter fouling in the dialyzer. To this end, we propose a three-in-one roller pump system to control the filtering/back-filtering conditions with the adjustment of flow-rate difference between the inlet and outlet of the dialysate. Furthermore, we investigated the flow rate owing to the pump-shell with/without a gutter. The experimental results revealed higher flow rates with a single gutter in the center of the shell. To expand this evidence for a deformable pump shell to control the flow rates, we prototyped a shell with movable wall and its flow rate variation in fixed and movable conditions. Experimental results show significantly higher flow rate with the movable condition. To utilize these facts, we propose a deformable pump shell for a low-cost, compact, and flow rate adjustable pump of the anti-fouling hemodialysis.

GS9-5 Development of a guidance and capture control system for a Post-tethered docking mechanism for micro-satellites

Yuto Hara[1] and Katsuyoshi Tsujita[2]
([1]Graduate School of Sustainability Science, Tottori University, Japan)
([2]Faculty of Engineering, Tottori University, Japan)

In recent years, the diversification of missions has necessitated the construction of large-scale space structures. In particular, there is a need for technologies to realize “in-orbit assembly,” where multiple satellites autonomously dock after orbital insertion. To address this technical requirement, we are researching to develop and demonstrate a new docking method centered on the post-tether docking technique, targeting docking between ultra-small satellites. In this study, we specifically developed part of a module intended for deploying the tether used in this method and conducted ground experiments to evaluate its characteristics. This module comprises a sensor system to estimate the position of the tether's tip, an actuator system to guide it to the desired position, a gripping system to grasp the component, and a computer that integrates these systems. Experimental results demonstrated that, within the range of assumptions made for the experimental setup, this module possesses sufficiently robust deployment capability and does not require complex calculations.

GS9-6 Study on Fault Prediction of Factory Equipment Using Fan Operating Sound

Asato Doi, Nobuo Iwasaki and Kazuya Okamoto
(National Institute of Technology, Wakayama College, Japan)

This study presents a low-cost, real-time fan fault detection system using Arduino-based acoustic analysis. Acoustic data from a DC fan were collected with the onboard PDM microphone of the Arduino Nano 33 BLE Sense and processed using time- and frequency-domain features. The normal operating sound exhibited stable and consistent spectral characteristics, forming a baseline reference for condition monitoring. Future work will include acquiring abnormal fan noise caused by airflow obstruction and other mechanical faults to build a robust classifier for distinguishing normal and faulty states. The ultimate goal is to develop an embedded and lightweight monitoring system that enables early detection of fan anomalies and supports predictive maintenance in practical applications.

January 21 (Wednesday), 16:05-17:35

B1 Meeting Room 2

GS13 Human-machine interaction and collaboration II

Chair: Yuya Okadome (Tokyo University of Science, Japan)

GS13-1 Wearable vs. World-Grounded Haptics for Space Teleoperation: A VR-Based Comparative Analysis of Kinesthetic Feedback for Fine Telemanipulation

Keshavi Hiren Joshi, Rene M. Suarez Flores, Muhammad Abdullah Zahid and Sajid Nisar
(Kyoto University of Advanced Science, Japan)

Space teleoperation systems are commonly operated using world-grounded interfaces that rely primarily on visual feedback, which can increase cognitive load and pose safety risks during fine manipulation tasks. To address these limitations, kinesthetic force feedback has been introduced as an additional sensory channel. Recently, wearable haptic devices have gained attention due to their potential for intuitive control and increased freedom of movement; however, systematic comparisons with traditional world-grounded haptic devices in space teleoperation contexts remain limited. In this work, we present a comparative evaluation of a world-grounded haptic device and a wearable haptic device within a virtual reality (VR) environment. A task designed in Unity to simulate representative motion and manipulation actions relevant to space teleoperation is used to assess performance under four conditions: wearable and world-grounded control, each with and without force feedback. Task accuracy and completion time are analyzed to evaluate the influence of device type and haptic feedback on operator performance. The results provide initial insights into the strengths and limitations of wearable and world-grounded haptic interfaces for fine telemanipulation tasks, highlighting their respective suitability for future space teleoperation systems.

GS13-2 Effects of Dimensionality Reduction Methods on Motor Learning in Controlling High-DoF Virtual Hands

Koki Onishi, Wen Liang Yeoh and Osamu Fukuda
(Saga University, Japan)

Direct control of complex, multi-DoF robotic devices is often impractical due to the high cognitive load required.While partial autonomy can reduce this burden, it may introduce coordination issues in applications requiring close human collaboration, such as prosthetics, due to the unpredictability of the autonomous actions. This highlights the need to maintain full user control while minimizing cognitive effort. Dimensionality reduction addresses this challenge by generating high-dimensional movements from minimal input signals, providing continuous control and flexibility, though it typically involves a steep learning curve. This study examines the comparative effects of the two common dimensionality reduction methods, Principal Component Analysis (PCA) and Non-negative Matrix Factorization (NNMF). Specifically, we investigate how the choice between PCA and NNMF influences a user's ability to control a virtual multi-DoF device over time. Four participants practiced using the device with PCA- and NNMF-based control schemes in ten 2.5-minute sessions conducted on separate days. The results demonstrate performance improvements over time for both methods and indicate that the NNMF-based control scheme is easier to learn and operate than the PCA-based scheme.

GS13-3 Design Element Decomposition of Assist Technologies Aimed at Inducing Flow States in Teleoperated Robot Operators: An Approach Based on Kansei Engineering

Takumi Kawamura
(Shunan University, Japan)

As the use of semi-autonomous teleoperated robots expands, achieving both collaborative efficiency and a positive psychological state for operators is a critical design challenge. However, a concrete methodology for robot design that satisfies both goals has not been established and explored feasible design methodologies to achieve a new design goal—“inducing operator flow state”—that could integrate these two objectives. This study comprehensively analyzed a dataset of metrics, including flow intensity, collected under a specific scenario to expand employment opportunities for individuals with severe motor dysfunction. It decomposed the highly abstract design goal into related, more concrete indirect design objectives, thereby structuring the design goals. Using the Kansei Engineering methodology, we performed multiple regression analysis on 117 trial samples from 11 participants (6 with severe motor dysfunction). Treating flow intensity as the objective variable and other metrics as design variables suggested that the following are critical elements as indirect design objectives: reducing downtime and temporal demand; increasing mental demand; challenge and skill level; and increasing translational operation. Furthermore, by positioning these elements as indirect design goals that mediate enhancements in flow intensity, we proposed multiple designs worthy of verification across two primary approaches and identified a design constraint.

GS13-4 A Detection and Evaluation Method for Finger Pointing and Calling Based on Element-Level Motion Classification for Multi-Workplace Applications

Kenta Yonejima and Akira Urashima
(Toyama Prefectural University, Japan)

Finger pointing and calling is a safety practice involving physical pointing and verbal confirmation to prevent occupational accidents and is widely adopted in Japanese industries. While its effectiveness in reducing human error has been demonstrated, providing individualized training remains challenging due to labor shortages. In our previous work, we developed a system to detect and evaluate finger pointing and calling in nursing workflows. However, the system was tailored to nursing-specific behaviors, making it challenging to apply to other fields such as the chemical industry. To address this issue, this study proposes a detection and evaluation method adaptable to various workplaces by focusing on fundamental behavioral elements that constitute finger pointing and calling. The proposed method defines a new “pointing” motion in addition to the “raising” and “lowering” motions. Using sensor data, these motions sequence, duration, and speech recognition results. Validation experiments using data from nursing and factory environments demonstrated that the proposed method outperformed conventional methods, indicating its applicability to different workplace environments.

GS13-5 Interactive Design and Evaluation Process for Ear-EEG SSVEP-BCI

Sodai Kondo and Hisaya Tanaka
(Kogakuin University, Japan)

The steady-state visual evoked potential (SSVEP) is a neural response with excellent information transmission capability, which makes it an attractive input source for brain-computer interfaces (BCI). The SSVEP is obtained from an electroencephalogram (EEG), which is typically measured on the scalp. Ear-EEG is more convenient and comfortable for the user because the ear is hairless, but the ear-EEG signal suffers from more attenuation than the conventional EEG signal. Designing an ear-EEG SSVEP-BCI and evaluating its performance is a long and extensive process owing to the need for long-term tests with multiple subjects. We previously developed a method for predicting the performance of an ear-EEG SSVEP-BCI by using a small number of trials but did not sufficiently explore its versatility and applicability. In this study, we developed an interactive BCI design and evaluation process using multiple methods and datasets. The results showed that the K-nearest neighbors and curve-fitting methods could both be used to successfully predict the BCI accuracy. In particular, we demonstrated that the BCI accuracy could be predicted even when the training and test data were independent of each other.

GS13-6 Investigating Cooperative Dynamics and Intention Inference in Asymmetric Physical Interaction

Hayato Eshita, Wen Liang Yeoh and Osamu Fukuda
(Saga University, Japan)

As technology advances, assistive robots providing physical support are becoming integral to daily life. These robots must physically interact with users for extended periods, exerting forces that augment human strength. Achieving such assistance requires precise coordination of intentions and timing; however, design methodologies for this remain unclear. This study aims to gain insights into improving human-robot coordination by investigating cooperation mechanisms in asymmetric relationships, modeled after real-world assistance scenarios. We hypothesized that repeated task execution would promote shared intent, thereby reducing operational errors and physical load. To test this, we conducted controlled experiments with human pairs performing a step-response task. Participants were assigned asymmetric roles: a User focused on the task and an Assistant focused on support. We analyzed how shared intent and coordination strategies evolved during these interactions. The experimental results did not support the hypothesis, as task repetition yielded no significant reduction in errors or load. However, these results provided valuable insights into the importance of temporal factors. The analysis indicates that establishing shared intent is challenging in short, one-off tasks. Consequently, the findings suggest that continuous, sufficiently long interactions, mirroring real-world support scenarios, are likely essential for achieving effective assistance and seamless cooperation.

January 21 (Wednesday), 16:05-17:20

B1 Meeting Room 3

GS20 Machine learning III

Chair: Tomohiro Harada (Saitama University, Japan)

GS20-1 Contrastive Pre-training of 3D Chest CT and Radiology Reports for Medical Image Diagnosis

Yuto Nomura[1], Satoru Ikebe[1], Shoji Kido[2] and Shingo Mabu[1]
([1]Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Japan)
([2]Institute for Radiation Science/Graduate School of Medicine, The University of Osaka, Japan)

Machine-based medical image diagnosis typically requires large amounts of task-specific annotations, which are costly and difficult to maintain. In clinical practice, however, radiology reports are routinely accumulated together with medical images, and effectively leveraging these textual descriptions may enhance diagnostic performance beyond image-only learning. Inspired by Contrastive Language–Image Pre-training (CLIP), which aligns images and text in a shared representation space, we propose a multimodal pretraining framework for chest CT analysis that learns the correspondence between images and radiology reports without using explicit diagnostic labels. In the pretraining stage, a 3D image encoder and a text encoder were trained using contrastive learning to align paired chest CT images and their associated reports in a shared embedding space. Normal and abnormal labels were not used during this process. For downstream evaluation, only the pretrained image encoder was used, with a multi-layer perceptron attached for binary classification (normal vs. abnormal). A baseline model without multimodal pretraining was constructed under identical conditions. Experimental results demonstrate that the proposed model consistently outperformed the baseline in both accuracy and AUC, while also exhibiting more balanced prediction behavior. These findings suggest that multimodal pretraining effectively enhances image representations and provides a promising foundation for medical image analysis.

GS20-2 Woven EndoBridge Device Size Prediction Using Machine Learning with Intracranial Aneurysm Morphological Features

Ayane Takahashi[1,2], Soichiro Fujimura[1,2], Genki Kudo[2], Shunsuke Hataoka[2], Gota Nagayama[2], Mitsuyoshi Watanabe[2], Yuichi Sasaki[2], Hiroyuki Enomoto[2], Hiroyuki Nishiyama[1], Hayato Ohwada[1], Toshihiro Ishibashi[2] and Yuichi Murayama[2]
([1]Tokyo University of Science, Japan)
([2]Jikei University School of Medicine, Japan)

Intracranial aneurysms (IAs) are cerebrovascular diseases characterized by localized dilation of cerebral arteries. The Woven EndoBridge (WEB) deployment procedure has been increasingly used as a minimally invasive treatment for bifurcation aneurysms; however, selecting the optimal device for individual cases while accounting for aneurysm morphology remains challenging. In this study, we developed a machine learning (ML)–based decision support system for WEB size selection by integrating patient clinical information with two- and three-dimensional morphological features of IAs to predict WEB width and height. Optimal WEB sizes were determined by clinicians through deployment experiments conducted in 273 bifurcation aneurysms, and device sizes were estimated from 31 candidate sizes using the proposed model. As a result, when the model presented a single size deemed most appropriate, the optimal WEB size actually used was included among the sizes recommended by the model in 47.3% of cases. When up to three candidate sizes were presented, the optimal WEB size was included in 76.4% of cases. These results demonstrate superior performance in size selection accuracy compared with conventional methods. Particularly under the single-candidate condition, the proposed approach outperformed clinician-based size selection. These findings indicated that the proposed ML system can objectively support WEB size selection.

GS20-3 Focal-SE MAE: Enhancing Spatial and Channel Dependency Modeling for Self-Supervised Feature Learning of Lung CT

Xilong Kang and Shingo Mabu
(YAMAGUCHI UNIVERSITY, Japan)

Accurate interpretation of lung CT scans is critical for early disease detection, yet standard Masked Autoencoders (MAE) often struggle to capture the subtle, low-contrast lesions typical of medical imaging. To address this limitation, this study proposes a novel self-supervised framework termed Focal-SE Masked Autoencoder (Focal-SE MAE). By integrating Focal Modulation and Squeeze-and-Excitation (SE) mechanisms into the transformer encoder, our method simultaneously enhances local spatial perception and inter-channel dependency modeling. Furthermore, a lightweight convolutional decoder is employed to maximize reconstruction fidelity while maintaining computational efficiency. We evaluated the model on the LUNA16 dataset using a 75% masking ratio. Experimental results demonstrate that Focal-SE MAE outperforms baseline MAE variants, achieving the lowest reconstruction error (MAE=0.0585) and the highest structural similarity (SSIM=0.902). These findings confirm that synergistically modeling spatial contexts and channel relationships yields superior feature representations, providing a robust foundation for automated lung disease analysis without relying on extensive manual annotations.

GS20-4 Improving Thyroid Cancer Classification by Generating Rare Case Images with a Diffusion Model

Hitoki Nakado[1], Shingo Mabu[1], Satoru Ikebe[1], Shohei Higuchi[2] and Kunihiro Inai[3]
([1]Yamaguchi University Graduate School of Science and Technology for Innovation, Japan)
([2]University of Fukui Hospital, Japan)
([3]Nagoya Institute of Technology Health Support Center, Japan)

Deep learning-based medical image analysis is often hindered by data scarcity and severe class imbalance, particularly for rare thyroid cancer subtypes. This study aims to address this issue by employing DreamBooth, a text-to-image diffusion model, to generate synthetic training data for classification. We utilized a dataset of HE-stained thyroid tissue images comprising seven classes, characterized by significant imbalance between common and rare carcinomas. By fine-tuning a Stable Diffusion model, we generated high-fidelity synthetic images that capture class-specific histological features. A ResNet50 classifier was trained on datasets combining real and synthetic images in varying proportions. The experiments revealed that datasets augmented with synthetic images yielded statistically significant improvements in classification accuracy compared to the baseline using real images only. Notably, adding synthetic data equivalent to 10% of the real dataset achieved the highest accuracy. Furthermore, F1-scores for rare classes, such as medullary and undifferentiated carcinomas, improved significantly. These results demonstrate that generative AI-based augmentation effectively mitigates class imbalance and enhances diagnostic performance in pathological image analysis.

GS20-5 Performance Evaluation of Anomaly Detection Model Configurations for Thyroid Tissue Images

Takuto Nishikawa[1], Shingo Mabu[1], Satoru Ikebe[1], Shohei Higuchi[2] and Kunihiro Inai[3]
([1]Yamaguchi University Graduate School of Science and Technology for Innovation, Japan)
([2]University of Fukui Hospital, Japan)
([3]Nagoya Institute of Technology Health Center, Japan)

The advancement of deep learning has accelerated its application in medical image analysis. However, disease detection tasks often suffer from data imbalance between classes due to varying data collection costs per class. Under such circumstances, anomaly detection methods are considered effective. While anomaly detection has been applied to thyroid tissue images, the exploration of optimal model configurations remains unexplored. This study aimed to clarify the characteristics of anomaly detection models suitable for thyroid tissue images by evaluating Autoencoder, DeepSVDD, and DeepSAD models under various configurations and experimental conditions. Specifically, we verified the performance of Autoencoder and DeepSVDD trained solely on normal images and examined the impact of preprocessing techniques such as masking, salt-and-pepper noise, and Gaussian noise. We also evaluated the performance of DeepSAD, which incorporates some abnormal images into training. The experimental results demonstrated that texture information is crucial for detecting abnormalities in thyroid tissue images. They also showed that preprocessing introducing disturbance on the image pixels contributes more to performance improvement than preprocessing that excludes some part of the pixels. Furthermore, we ontained insights regarding the required amount of abnormal data based on the relationship between AUC and data collection cost.

January 21 (Wednesday), 16:05-17:05

B1 Meeting Room 4

GS26 Mobile robots II

Chair: Kiyohiko Hattori (Tokyo Denki University, Japan)

GS26-1 Research and Development of Small Linked Crawler Robots for Search and Rescue Operations

Shunsuke Tamura[1], Jehun Seo[2] and Yoshiaki Yamazaki[3]
([1]Meisei University, Japan)
([2]Meisei University Graduate School of Science and Engineering, South Korea)
([3]Meisei University Graduate School of Science and Engineering, Japan)

This paper aims to develop a small crawler robot capable of locating people in need of rescue in areas where search and rescue dogs, rescue teams, and existing rescue robots have difficulty searching, such as narrow, uneven rubble or inside pipes. Currently available crawler robots often have a fixed structure with sub-crawlers, making them unable to move when they tip over. Therefore, I have begun developing a crawler robot with crawlers on all four sides, enabling it to move even when it tips over by using the crawlers in contact with the ground.

GS26-2 Design and Evaluation of a Variable Ground-Contact-Area Mechanism for a Legged Type Planetary Exploration Rover

Kensuke Mori[1], Kazuki Maeda[2] and Katsuyoshi Tsujita[3]
([1]Tottori graduate school, Japan)
([2]Tottori graduate school, Japan)
([3]Tottori University, Japan)

Autonomous planetary exploration demands high mobility over diverse environments, ranging from rigid rocky surfaces to soft regolith. To overcome the limitations of wheeled rovers, including slippage and sinkage, and the structural complexity of hybrid systems, we developed a "variable contact area mechanism" for legged robots. This system optimizes traversability by actively switching between point contact and surface contact modes depending on the terrain. This paper details the development of a quadruped prototype and reports on walking experiments conducted under varying conditions, including different gait periods and foot angles. Results confirm that the surface contact mode significantly improves travel distance and efficiency on soft ground by mitigating sinkage, while the point contact mode remains superior on hard surfaces, validating the effectiveness of the proposed adaptive mechanism.

GS26-3 Development of a Guarded Drone and its Autonomous Flight Algorithm in Indoor Narrow Space

Mitsuki Uesugi and Jae Hoon Lee
(Graduate School of Science and Engineering, Ehime University, Japan)

In recent years, the deterioration of infrastructure has increased the demand for efficient and safe inspection methods. Although drones are promising for structural inspection, indoor environments contain numerous obstacles and lack GPS, making stable autonomous flight difficult. To address this, this study develops a spherical protective guard and a LiDAR-based obstacle avoidance algorithm for safe flight in narrow, GPS-denied spaces. A lightweight CFRP spherical guard was developed to fully enclose the drone, allowing stable flight even upon contact with walls. A 360° LiDAR sensor was integrated for point-cloud-based SLAM localization and real-time path correction. Prototype tests showed that the guard prevented propeller collisions and maintained stable flight in a corridor of 1.2m width. Gazebo simulations further confirmed that the proposed algorithm generated smooth, obstacle-avoiding trajectories in narrow and complex environments. These findings indicate the potential of the proposed system for safe and efficient indoor inspections. Future work will extend the method to fully three-dimensional path planning toward practical autonomous inspection operations.

GS26-4 Wall-Following Exploration Flight Algorithm Using LiDAR in GPS-Denied Unknown Environments

Yudai Nagata and Jae Hoon Lee
(Graduate School of Science and Engineering, Ehime University, Japan)

In GPS-denied indoor environments, autonomous exploration requires accurate structure extraction and safe real-time path planning. This study proposes a wall-following algorithm that builds a 2D occupancy grid from LiDAR-based 3D point clouds using OctoMap, extracts wall boundaries as free cells adjacent to occupied or unknown cells, and forms continuous boundary loops. For each loop, offset points with a predefined safety margin are computed, and a smooth trajectory with uniform waypoint spacing is generated by B-spline interpolation and distance-based resampling. The exploration path is dynamically updated whenever newly observed regions modify the OctoMap, maintaining consistency with the latest environment map. Experiments in two indoor environments—one fully visible from the start and one with initially unknown regions—confirmed stable boundary extraction, path generation, and online path updates, as well as accurate tracking of the generated trajectories by a drone. These results show that the proposed algorithm provides a practical basis for autonomous indoor exploration and inspection.

January 22 (Thursday), 09:00-10:45

3F Meeting Room 32

OS25 ISBC: Individuality and Collectivity in Living Systems III

Organizers: Yasuhiro Hashimoto (The University of Aizu, Japan)
Akiko Kashiwagi (Hirosaki University, Japan)

OS25-1 Density-dependent differentiation and emergence of structured heterogeneity in clonal Tetrahymena thermophila

Akiko Kashiwagi[1], Hiroki Kojima[2] and Takashi Ikegami[2]
([1]Hirosaki University, Japan)
([2]The University of Tokyo, Japan)

Single-cell RNA sequencing (scRNA-seq) frequently reveals multiple transcriptional clusters even within genetically uniform (clonal) populations. This phenomenon has been reported across various microorganisms that proliferate by binary fission, indicating that genetically identical cells can diverge into distinct transcriptional states. Such diversification raises the possibility that population-level context contributes to the emergence of individuality within clonal populations; however, the origin and regulation of these heterogeneous states remain largely unexplored.
In our previous scRNA-seq analysis of Tetrahymena thermophila, we identified six transcriptional clusters on a UMAP projection, reflecting heterogeneous gene-expression states within a clonal population. To examine whether population density contributes to this diversity, we experimentally manipulated cell density and analyzed transcriptional changes using reverse transcription quantitative PCR (RT-qPCR).
Our results suggest that population density modulates transcriptional variability even under logarithmic growth conditions, where cells are generally assumed to share a uniform internal state, by reshaping the underlying transcriptional landscape and biasing how single-cell heterogeneity is expressed within clonal populations.

OS25-2 Information flow from macro to micro: a case study in insect collectives

Shigeto Dobata[1], Yusuke Notomi[1], Tatsumi Kudo[1], Ilya Horiguchi[1], Michael Crosscombe[1], Norihiro Maruyama[1], Acer Yu-Chan Chang[2] and Takashi Ikegami[1]
([1]The University of Tokyo, Japan)
([2]Helsinki Collegium for Advanced Studies, Helsinki Institute of Life Science, Finland)

We propose an information-theoretic framework to quantify directed “macro-to-micro” influence in collective behavior of social insects. Using transfer entropy and partial information decomposition, we separate unique, redundant, and synergistic contributions of an individual microstate and a coarse-grained macrostate to future individual activity across time lags. Applied to ant (Pristomyrmex punctatus) aggregation data (cluster membership as microstate; proportion of wanderers as macrostate), the method reveals substantial redundancy at short lags and little synergy. Some individuals exhibit delayed unique information from the macrostate, suggesting slow indirect feedback channels shaping collective aggregation dynamics.

OS25-3 Behavioral Individuality in Ant Collectives as Differential Usage of a Shared Behavioral Mechanism

Norihiro Maruyama, Michael Crosscombe, Shigeto Dobata and Takashi Ikegami
(The University of Tokyo, Japan)

Previous studies revealed that ants exhibit qualitatively different movement patterns—deterministic exploration and stochastic clustering—which can be captured by individual-specific ϵ-transducers reconstructed from behavioral time series. In this study, we extend this framework by constructing a single shared ϵ-transducer that represents the behavioral dynamics of all ants in the population. Instead of reconstructing an automaton for each individual, we infer a single finite-state machine from all observed trajectories across individuals and time. We then characterize individual behavior by identifying which states and transitions of this automaton are actually utilized by each ant. Our results suggest that, while the underlying behavioral structure is shared, individual differences emerge as selective usage of distinct part of the automaton. Finally, we perform simulations driven by the reconstructed ϵ-transducer and empirically estimated parameters. Although large-scale aggregation is difficult to reproduce, the simulations generate small clusters consisting of several individuals. The aim of this study is that individuality in ant behavior can be understood as constrained navigation within a shared behavioral grammar, rather than as fundamentally distinct internal mechanisms.

OS25-4 Exploring Cultural Evolution Through Modular Dynamics in Temporal Hashtag Networks

Yasuhiro Hashimoto[1], Hiroki Sato[2] and Takashi Ikegami[2]
([1]The University of Aizu, Japan)
([2]The University of Tokyo, Japan)

Biological, cultural and technological evolution have never stopped production of novelties. This phenomenon is called Open-Ended Evolution (OEE). This study analyzes the expansion of novelty on social media hashtags to investigate OEE mechanisms. Our empirical analysis identifies two distinct topological structures in novelty expansion in the evolutionary space of tag semantics: dense clusters termed "bunches" driven by popularity, and elongated chains termed "branches" driven by successive inheritance. We also modeled tagging behaviour replicating these structures of novelty production. The model and estimated parameter values form the empirical data suggested the novelty expansion in the empirical system cannot be explained solely by social effect (preferential attachment) and importance of successive inheritance that extend the space of possible novelties.

OS25-5 Quantifying and Modeling Novelty Production by Collective Creativity on the Web

Hiroki Sato[1], Yasuhiro Hashimoto[2] and Takashi Ikegami[3]
([1]The University of Tokyo/Alternative Machine Inc., Japan)
([2]University of Aizu, Japan)
([3]The University of Tokyo, Japan)

Biological, cultural and technological evolution have never stopped production of novelties. This phenomenon is called Open-Ended Evolution (OEE). This study analyzes the expansion of novelty on social media hashtags to investigate OEE mechanisms. Our empirical analysis identifies two distinct topological structures in novelty expansion in the evolutionary space of tag semantics: dense clusters termed "bunches" driven by popularity, and elongated chains termed "branches" driven by successive inheritance. We also modeled tagging behaviour replicating these structures of novelty production. The model and estimated parameter values form the empirical data suggested the novelty expansion in the empirical system cannot be explained solely by social effect (preferential attachment) and importance of successive inheritance that extend the space of possible novelties.

Invited Talk 1 Community First Theory: How Collective Organization Generates Individual Diversity Across Biological, Robotic, and LLM Systems

Takashi Ikegami (The University of Tokyo, Japan)

See page 14

January 22 (Thursday), 09:00-10:45

3F Meeting Room 33

GS7 Bio-inspired robotics

Chair: Norimitsu Sakagami (Ryukoku University, Japan)

GS7-1 Development of a lightweight and flexible tensegrity manipulator for mounting on mobile platforms

Kenshin Shimohara and Shuhei Ikemoto
(Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan)

Tensegrity-based continuum manipulators have attracted increasing attention due to their high strength-to-weight ratio, structural compliance, and suitability for safe interaction with the environment. These characteristics make them promising candidates for integration with mobile platforms such as drones or ground robots. In this study, we develop a lightweight, tower-shaped Class-1 tensegrity manipulator composed of five serially connected tensegrity prisms, each consisting of three rigid struts and tensioned cables. The manipulator is actuated by six compact servo motors driving selected cables, while elastic elements are incorporated to provide passive restoring forces for upright recovery. We propose a hybrid control strategy combining angle control and current-based limitation to achieve smooth bending motions while preventing excessive motor load. The motion performance and repeatability of the manipulator tip are quantitatively evaluated using a motion capture system through repeated bending experiments in multiple directions. Experimental results demonstrate stable bending behavior and consistent end-effector positioning, confirming the effectiveness of the proposed mechanical design and control approach. This work contributes to the development of lightweight, compliant continuum manipulators suitable for deployment on mobile robotic platforms.

GS7-2 A Low-Cost Three-Layer Sensing Architecture for Reflexive Grip Control in Prosthetic and Robotic Hands

Shubh Shukla and Sajid Nisar
(Kyoto University of Advanced Science, Japan)

Accurate grip control in prosthetic hands usually demands expensive tactile arrays or multi-layer sensor fusion, making practical systems unaffordable for cost-sensitive applications. To mitigate this limitation, we propose a low-cost three-layer sensing system that realises stable and adaptive grasping under the restrictions of off-the-shelf appearance only. The first layer has a simple micro-switch to monitor fully open or closed positions in order to protect the motor drive and provide mechanical reference point. Based on motor-current feedback; in this way approximate force measurement can be realized after calibration without special force sensors. In the third layer, a piezoelectric vibration sensor at the fingertip is used to detect incipient slip and generate a rapid reflex-like increase of grip force as seen in human tactile reaction. The main contribution is the amalgamation of these two sensing modalities into a cohesive biomimetic framework that addresses the application requirement for mechanical safety and dynamic grasp stabilization. This type of detailed approach with current-based force estimation as well as vibration-based slip detection has not previously been realized before in prosthetic hand systems. Bench tests show that slip can be detected in ~20 ms and force estimation is still within about ±10 percent of calibration levels. These results demonstrate that a high-order level of control over grasping can be obtained with inexpensive devices and low computational requirements, and that this approach is suitable for real-world prosthesis as well as assistive robotics applications.

GS7-3 Dynamical Effect of Passive Anatomy Trains for Musculoskeletal Humanoid Robot Model

Kazuaki Yaegashi[1], Soma Kato[1], Shunsuke Fujii[1], Tomoya Kamimura[2] and Akihito Sano[1]
([1]Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology, Japan)
([2]Department of Mechanical Science and Bioengineering,
Graduate School of Engineering and Science, The University of Osaka)

Human locomotion is produced by the dynamic interaction between the whole body and the environment, and the significance of these interactions has recently been recognized in the field of machine learning. In this study, we focus on interactions between the passive body and the Anatomy Trains (AT), based on anatomical insights, especially the Spiral Line (SPL), a double-helix fascial pathway thought to support walking stability. We construct musculoskeletal humanoid models with a passive upper body with and without SPL passive pathways. Results of learning with identical reward functions and hyperparameters using deep reinforcement learning, without SPL models, fail to achieve forward locomotion, whereas models with SPL achieve continuous forward walking, with SPL generating upper-body torsion that supports leg swing. These results suggest that the interaction between the SPL pathway and a passive body achieves the generation and stabilization of walking performance, and these findings may contribute to biomechanical analysis of human locomotion.

GS7-4 Development of a Biomimetic Robot Hand Faithful to Bone Geometry

Haruto Kimata, Satoshi Suenaga and Shuhei Ikemoto
(Kyushu Institute of Technology, Japan)

Biomimetics is one of the promising approaches for reproducing the human hand by robotics. However, due to its structural complexity, comprehensive imitation remains challenging, and a five-fingered biomimetic robot hand applicable to real tasks has yet to be fully realized. Therefore, this study aims for faithful reproduction of bones' geometry based on a human hand skeleton model while adopting an engineering simplification for muscle, tendon, and ligament by a silicone rubber membrane. This simplification lacks biological faithfulness but defines relative positions and motions of bones similar to them of the human hand. The membrane combines a flexible silicone sheet and a stiff resin film to achieve both the reproduction of flexibility allowing joint motion and the strength necessary for joint stability. This paper describes the design and fabrication procedures of the developed biomimetic robot hand in detail. Through evaluation experiments of the developed hand, the realization of thumb opposition and sufficient grasp stability based on quantitative evaluation were confirmed. These results demonstrate the validity of the proposed approach.

GS7-5 Design and evaluation of artificial muscles drive system by an artificial spinal cord IC for generating gait patterns

Yugo Kokubun, Tasuku Hirasawa, Ryukichi Nakamura, Ken Saito, Fumio Uchikoba and Minami Kaneko
(Nihon University, Japan)

Most humanoid robots drive each joint with servo motors, controlling joint angles through massive numerical calculations performed by CPUs and software. In contrast, humans efficiently generate signals to control muscle contraction and relaxation via neural networks in the brain and spinal cord. We are developing musculoskeletal humanoid robots that mimic human movement generation mechanisms and musculoskeletal structures. Based on physiological findings, we have constructed an artificial spinal cord circuit that mimics the spinal cord's CPG using analog electronic circuits. We have verified neural signal patterns during walking and running using a circuit simulator. Furthermore, based on gait dynamics analysis, we developed 12 types of artificial muscles using shape memory alloys and elastomers, and verified joint movements using a skeletal model. This paper reports the implementation of the artificial spinal cord circuit onto a 2.4 mm square IC and the verification of its gait pattern output characteristics. Furthermore, an amplification and conversion circuit was constructed to drive the artificial muscles according to the IC output timing. Using LTspice®, it was confirmed that the desired current flows to each artificial muscle at the two timing points during walking when each muscle generates its maximum force.

GS7-6 Ladybug-Inspired Flapping Motion: Kinematic Model and Mechanism

Kazuki Takeshita[1], Geunho Lee[1], Kosei Shiinoki[1], Kota Okabe[2] and Min Kee Park[3]
([1]University of Miyazaki, Japan)
([2]Miyazaki Airport Building CO. LTD., Japan)
([3]Seoul National University of Science & Technology, South Korea)

Rescue robotics has evolved, integrating various mobility mechanisms to traverse debris and overcome barriers. However, reliance solely on ground-based systems presents significant limitations in navigating complex disaster environments. To overcome these limitations, aerial mobility mechanisms have gained interest for their potential to enable robotic traversal and flight in complex terrain. In addition to conventional aerial mechanisms such as airplanes and helicopters, biologically inspired flapping mechanisms that mimic the motions of birds and insects have also been investigated. In line with this trend, this study investigates the ladybug-inspired flapping flight mechanism. Initially, the flapping motions of ladybugs are systematically observed and analyzed to extract their key kinematic characteristics. Based on these observations, a kinematic flapping model is developed. To assess the model's capability to generate aerodynamic lift, simulations and evaluation experiments are performed. Finally, the effectiveness of the proposed flapping model is validated, and implications for future research are discussed.

GS7-7 Smart Toilet Seat for Automatic Cleaning: Design and Implementation

Kazuki Takeshita[1], Geunho Lee[1], Kosei Shiinoki[1] and Kota Okabe[2]
([1]University of Miyazaki, Japan)
([2]Miyazaki Airport Building CO. LTD., Japan)

During the pandemic of Covid-19, there has been a remarkable increase in global interest in a wide range of infections. The etiology of these diseases is typically attributed to the direct and indirect infections through contact with biological fluids containing pathogens, such as viruses and bacteria. This method of infection poses an exceptionally high risk in places where a significant number of unidentified people gather. Notably, concerns have been raised regarding the sanitation of their toilets due to the toilet seats, which has a potential for facilitating indirect infection. To prevent these infections, this study presents a novel type of smart toilet seat designed for automatic cleaning. First, as a fundamentally crucial concept, the timing for initiating the cleaning process is considered to complete the operation while preserving the convenience of toilet users. Next, the belt-pulley-type prototype is developed, and the realization of the proposed cycle is confirmed. Furthermore, the developed prototype is enhanced by the incorporation of a flexible rack. Using this prototype, extensive simulations and experiments are demonstrated, validating the effectiveness of the proposed smart toilet seat. In the future, the implementation in public restrooms will contribute to the maintenance of the public health.

January 22 (Thursday), 09:00-10:30

B1 Meeting Room 1

GS10 Data mining

Chair: Shudai Ishikawa (National Institute of Technology, Oita College, Japan)

GS10-1 Causal Discovery in Additive Noise Models using Beam Search

Hans Jarett Ong, Brian Godwin Lim, Renzo Roel Tan and Kazushi Ikeda
(Nara Institute of Science and Technology, Japan)

Causal discovery from observational data is a fundamental challenge. Greedy search algorithms like Regression with Subsequent Independence Test (RESIT) are susceptible to making irreversible errors, especially in high-variance contexts caused by unmeasured confounders or finite samples. To address this, we introduce a novel generalization of RESIT that replaces its local, greedy search with a robust beam search, framing the task as a path search on a state-space graph. Through extensive simulations, we demonstrate that structural accuracy (SHD, SID) consistently improves as the beam width increases. We find these gains are most statistically significant in intermediate sample-size regimes (n=250, 500), where statistical noise is high enough to mislead the greedy search, an error our wider beam search corrects. Our framework provides a practical, tunable algorithm that bridges the gap between fast but brittle local searches and computationally infeasible global searches, enhancing causal discovery in complex settings.

GS10-2 Estimation of Travel Mode through Unsupervised Learning Using GPS data and GIS

Misaki Fukui[1] and Hiroyasu Inoue[1,2]
([1]Graduate School of Information Science, University of Hyogo, Japan)
([2]Center for Computational Science, RIKEN, Japan)

Understanding actual travel behavior is essential for effective urban planning, yet traditional travel surveys are often costly and lack real-time coverage. This study proposes an unsupervised travel mode estimation method that integrates Global Positioning System (GPS) trajectory data with Geographic Information System (GIS) information. Using data collected in Nagasaki Prefecture, we segmented trajectories and extracted basic motion features—such as speed, acceleration, and direction change rate—alongside GIS-based proximity rates to railway and bus networks. Travel modes were estimated using Fuzzy c-means clustering to account for the ambiguity inherent in human movement. The identified clusters were classified as train, bus, car, and two-wheeler. When comparing the estimated modal shares with the National Person Trip Survey, results indicated that train and two-wheeler modes were estimated with reasonable consistency. However, significant misclassification occurred between car and bus modes due to their similar kinematic characteristics. Notably, when these modes were aggregated into a single “motorized road transport” category, the estimated shares aligned closely with survey data. These findings demonstrate that the proposed method can capture realistic travel behavior at a coarse classification level.

GS10-3 Predicting Patient Volume Anomalies in Outpatient Phlebotomy Units: A Multi-Step Ahead Forecasting Framework

Tzu-Chi Liu[1], Chun Chang[1], Ming-Shu Chen[2], Chih-Te Yang[3], I-Fei Chen[3], Chien-Chih Wang[4] and Chi-Jie Lu[1]
([1]Fu Jen Catholic University, Taiwan)
([2]Asia Eastern University of Science and Technology, Taiwan)
([3]Tamkang University, Taiwan)
([4]Ming Chi University of Technology, Taiwan)

The rapid rise in outpatient visits has made outpatient phlebotomy units (OPUs) a major bottleneck, where unpredictable blood-test demand and limited specialized staffing cause overcrowding, long waits, and reduced patient satisfaction and safety. Proactive capacity management is essential because staffing adjustments require significant lead time. To support early intervention, this study develops a multi-stage ensemble temporal value multi-step ahead forecasting (MSETVMF) model to predict patient-volume anomalies that exceed workforce capacity. The model enriches traditional direct and recursive forecasting by aggregating temporal features from three machine learning models—XGBoost, Random Forest, and LightGBM. Using real OPU data from a Taiwanese medical center, the model showed strong performance for 2- to 4-step ahead forecasting, with Random Forest best at 2- and 4-step and LightGBM at 3-step. The approach offers a practical tool for proactive resource allocation.

GS10-4 An Additive Spline Classification Model by Minimizing Derivatives

Tatsumasa Sakai, Yoshifumi Kusunoki and Tomoharu Nakashima
(Osaka Metropolitan University, Japan)

This paper addresses the additive spline classification model, which assigns class labels based on the sum of feature-wise spline functions. To enhance the generalization performance of this model through margin maximization, we propose a novel regularization method that minimizes the derivatives of spline functions with respect to input features. Specifically, minimizing first-order derivatives reduces the steepness of the function slopes to enlarge the classification margin, while minimizing higher-order derivatives promotes function smoothness. The proposed regularization is formulated as a convex quadratic function, facilitating efficient optimization using existing software. We evaluate the performance of the proposed method through numerical experiments on a two-dimensional synthetic data and benchmark datasets, to compare it with conventional regularization that directly minimizes spline parameters. We analyze the effects on classification boundaries, hyperparameter sensitivity, and feature-wise explainability. The results demonstrate that the derivative-based regularization significantly improves both classification accuracy and the interpretability of the trained spline functions.

GS10-5 Research on the sustainable production inventory model based on data-driven and progressive carbon tax policy

Ruey-Chyn Tsaur[1], Ya Hsuan Wang[1], Nei-Chih Lin[1], Chi-Jie Lu[2], Yin-Yin Huang[3] and Chih-Te Yang[1]
([1]Tamkang University, Taiwan)
([2]Fu Jen Catholic University, Taiwan)
([3]Nanchang Vocational University, China)

The increasing pressure of climate change and digital transformation has intensified the need for sustainable and data-driven production–inventory management in supply chains. This study develops a data-driven sustainable production-inventory model for a two-echelon supply chain that integrates demand forecasting, low-carbon material selection, carbon-reduction investment, and progressive carbon tax policies. A two-stage decision framework is adopted: in the first stage, market demand is estimated using time-series forecasting methods based on historical data; in the second stage, an integrated production-inventory optimization model is solved to determine optimal procurement, production, shipment, replenishment, and carbon-reduction investment decisions under a progressive carbon tax scheme. Numerical experiments demonstrate that incorporating demand forecasting improves decision accuracy and overall profitability, while progressive carbon taxes effectively incentivize emission reduction through operational and technological adjustments. Sensitivity analysis on the proportion of low-carbon raw materials reveals a clear trade-off between emission reduction and total integrated profit, indicating that moderate adoption levels can achieve a better balance between economic performance and environmental sustainability. The proposed model provides practical managerial insights for aligning supply-chain operations with carbon reduction policies and long-term sustainability objectives.

GS10-6 How Social Media Shapes Public Oversight of Law Enforcement in Indonesia

Muhammad Asri Safi'Ie[1], Kunhao Yang[2], Nurul Firdaus[3], Satoru Ikebe[1] and Shingo Mabu[1]
([1]Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Japan)
([2]School of Engineering, Shibaura Institute of Technology, Japan)
([3]Informatics Engineering, Sebelas Maret University, Indonesia)

Public trust in Indonesian law enforcement has significantly declined over the past decade, driven by allegations of official misconduct and a perceived lack of transparency. The high-profile murder case of Brigadier Joshua by his superior, former Inspector General Ferdy Sambo, ignited widespread public outcry and highlighted the emerging role of social media as a crucial tool for social control and legal oversight in Indonesia. This study investigated how social media and the public leveraged platforms to pressure legal institutions during the Ferdy Sambo case. YouTube data, comprising video descriptions and netizens' comments, were collected using the keyword "Ferdy Sambo." The data were analyzed using heatmap analysis, Gini coefficient, grouped boxplots, and grouped barplots. Sentiment analysis measured collective public emotion, while engagement analysis examined how digital media and audiences influenced law enforcement actions. The study found that digital media reporting significantly influenced public awareness, quickly turning news into viral topics. Public suspicion of legal manipulation was immediate, fueled by the involvement of officers in obstruction of justice and a fabricated crime scene. Digital media fostered greater transparency, and amplified public commentary generated calls for accountability and applied significant pressure for justice. Online commentary was highly punitive. Results showed that cases receiving widespread, viral attention were processed more rapidly compared to less visible cases. This study demonstrates the potent ability of social media to mobilize collective awareness and exert public pressure on Indonesia's legal institutions. This collective online oversight contributes to increased transparency and faster legal processing in high-profile cases, underscoring social media's critical role in demanding justice and accountability.

January 22 (Thursday), 09:00-10:45

B1 Meeting Room 2

GS14 Human-machine interaction and collaboration III

Chair: TBD

GS14-1 HARVEE: Research on AI-based Hybrid System Generating Cooperative Behavior between Human and Abstract Object

Ayano Yasui
(Institute of Innovation, Musashino Art University, Japan)

In this study, we aimed to realize cooperative interaction between a human and a physical agent. Guided by a design philosophy that explores interaction through an artistic, abstract object devoid of biological cues, we developed HARVEE, a physical agent system built upon our previous research. The agent uses a physical body composed of an active pendulum mechanism driven by two servo motors and an internal spring structure to respond to human movements in real time, thereby generating a cooperative feedback loop between the human and the agent. A feature of the system's software algorithm is to generate diverse and non-predefined behaviors by determining a characteristic of human movement from skeleton data detected by recognition AI, and smoothly fusing multiple motion modes according to the person's movements through weighting. An experiment involving seven participants demonstrated that the system, by continuously changing its diverse motions, including both autonomous and tracking responses, was able to elicit sustained and exploratory behavior from users, leading to the establishment of cooperative interaction even with minimal agent movement.

GS14-2 Improvement of the Pressing Cover of the Guitar Performance Assistive Device "F-Ready" for Individuals with Cerebral Palsy

Aki Yamauchi, Yukiko Nishinohira and Hirokazu Matsui
(Department of Mechanical Engineering, University of Mie, Japan)

This study examines grip and lever design approaches that may facilitate more stable operation of hand-controlled devices for individuals with cerebral palsy, including those with severe motor impairments. People with cerebral palsy often experience abnormal muscle tone and limited motor coordination, which can make grasping and lever manipulation difficult, particularly when wrist angles vary. To address these challenges, a grip shape is proposed that is intended to support grasping using the fingers and palm, as well as finger-based lever-pulling motions under different wrist postures. Several grip designs with different lever configurations were developed and evaluated through experiments involving participants with cerebral palsy. Usability was assessed based on subjective evaluations of operability and qualitative observations of hand posture and finger usage. The results suggest that grip and lever design differences may influence operational stability, and that suitable design parameters for individuals with cerebral palsy may exist within a relatively limited range. Furthermore, the proposed grip was applied to the grip section of the guitar performance support device F-Ready, developed by our research group, and its practical usability was examined through actual use.

GS14-3 Robot Automated Design for Realizing Personalized Robots

Katsuma Akamatsu, Sho Yamauchi and Keiji Suzuki
(Future University Hakodate, Japan)

Personal communication robots have gained attention in recent years. However, their widespread adoption remains limited due to difficulties in fostering long-term user attachment and the lack of diversity in robot appearance and behavior. Allowing users to participate in robot design has been suggested as an effective approach, but existing methods often require expert knowledge. This study proposes an end-to-end robot creation system that enables users to design customized robots from simple hand-drawn sketches. The system generates a 3D appearance model from a sketch, automatically estimates skeleton structures, and produces multiple motion patterns using a diffusion-based motion generation model. The generated motions are converted into servo motor control signals and reproduced on a physical robot using the proposed system. Experimental results demonstrate that diverse robot appearances and motions can be generated without specialized skills and executed on real robot hardware. These findings demonstrated the potential of the proposed system as a framework for personalized robot creation, while identifying several challenges.

GS14-4 Design and Development of a Compact Tendon-Driven Soft Robotic Glove for Hand Rehabilitation

Queena Qurrota Ayun, Khalid Meitani and Sajid Nisar
(Kyoto University of Advanced Science, Japan)

Hand motor impairments resulting from neurological and musculoskeletal conditions often require long-term, repetitive rehabilitation to restore functional movement. Conventional therapy sessions require frequent clinical visits and bulky devices, eliminating the accessibility of home-based therapy. Therefore, this paper presents a 3D printable compact tendon-driven soft wearable robotic glove fabricated from thermoplastic polyurethane (TPU) intended for home-based rehabilitation. To minimize distal mass and improve comfort, all actuators are integrated into a wrist-mounted unit, reducing the load on the hand while maintaining effective finger actuation. Unlike many existing designs that rely on distributed actuators or sensors placed on the hand, the proposed glove prioritizes mass relocation and material compliance to enhance wearability. The soft robotic glove prototype weighs 184 g in the upper limb, resulting in the lowest weight among previous works of five-finger actuation gloves. Experiments were conducted with seven healthy subjects who were instructed to keep their hands stationary throughout the trials, and the results showed generally positive subjective evaluations in terms of comfort, grasping motion quality, perceived assistance, and perceived safety.

GS14-5 A preliminary study to evaluate the impressions of SPRO (Selfish Pretty RObot), an inconvenient robot that promotes behavioral change

Miyu Fujisawa[1], Takehito Kikuchi[2] and Kenichiro Tanaka[3]
([1]Graduate School of Science and Technology, Oita University, Japan)
([2]Faculty of Science and Technology, Oita University, Japan)
([3]Faculty of Welfare and Health Sciences, Oita University, Japan)

Population aging is rapidly progressing worldwide, with the proportion of older adults increasing in nearly every country. Japan, in particular, represents the forefront of this demographic shift, where maintaining physical and cognitive health and extending healthy life expectancy have become urgent issues. To address this issue, we have been developing the Selfish Pretty RObot (SPRO), a robot deliberately designed with “inconvenience,” with the aim of providing older adults with a sense of daily meaning in life. SPRO is intended to function as a substitute for companion animals, offering opportunities to go out and thereby contributing to the maintenance of physical and psychological health. The system consists of three elements: the SPRO robot itself, a dedicated charger called “ECSPOT,” and a smartphone application. In addition to increasing outdoor frequency, the system is expected to alleviate loneliness and enhance psychological fulfillment. This article reports on the development of the first prototype of SPRO and presents the results of a preliminary hearing survey conducted with 13 older adults aged 70 and older in Kitsuki City, Oita Prefecture. The findings reveal several key design challenges, including the robot’s shape, aesthetic design, and sound volume, highlighting important directions for improvement.

GS14-6 Enhancing Task Efficiency of Myoelectric Prosthetic Hands Through Vision-Based Object Tracking

Xiqing Di, Ahmad Mowaz, Nana Fukuta, Wen Liang Yeoh and Osamu Fukuda
(Saga University, Japan)

With the rapid advancement of computer vision technology, incorporating vision sensors into conventional myoelectric prosthetic hands presents a promising approach to enhancing both functionality and user experience. This research proposes a Dual-sensor control system that integrates electromyography (EMG) sensors with a vision sensor. The system consists of two primary modules: an EMG signal recognition module, based on the Log-Linearized Gaussian Mixture Network (LLGMN), and a vision module utilizing an AI camera for real-time object tracking and orientation adjustment. Based on this system, we developed a vision-based myoelectric hand using AI cameras. The vision-based myoelectric hand is capable of performing four distinct movements, effectively simulating the process of object grasping. To validate the feasibility and effectiveness of this system, we conducted experiments on picking tasks. Experimental results demonstrate that, compared to traditional EMG-only control, the integration of the vision sensor significantly reduces task completion time. These results confirm the feasibility and effectiveness of combining EMG and computer vision, and lay the groundwork for future applications involving more advanced visual processing techniques.

GS14-7 Evaluation of a Solution Support System for 2D Assembly Tasks

Ryosuke Yamamoto, Wen Liang Yeoh and Osamu Fukuda
(Saga University, Japan)

In manufacturing assembly processes, the way instructions are presented is thought to affect productivity and workers' mental burden. Particularly when tasks are complex or involve multiple steps, understanding the work becomes difficult, leading to decreased productivity and increased mental strain. Therefore, this study investigates how differences in the instructions given to workers during two-dimensional assembly tasks affect productivity and mental burden. We adopted silhouette-puzzle problems as the assembly task, developed an assembly support system using image recognition, and measured completion time and subjective evaluations. The proposed system supports cognitive functions, such as judgment and selection, required for solving silhouette puzzle problems. Ten young adults (age: 22.0 ±1.0 years) were instructed to reproduce shapes displayed on a screen using tetromino blocks. The verification experiment revealed that while the system's thinking support did not significantly reduce measured time, the interactivity of the instructions potentially alleviated mental burden.

January 22 (Thursday), 09:00-10:30

B1 Meeting Room 3

GS23 Machine learning VI & Neural networks

Chair: Sutthiphong Srigrarom (National University of Singapore, Singapore)

GS23-1 A Study on the Explainability of SIRMs Fuzzy Classifiers with Data-Driven Module Structures

Yuki Yamasaki, Yoshifumi Kusunoki and Tomoharu Nakashima
(Osaka Metropolitan University, Japan)

Single Input Rule Modules (SIRMs) fuzzy classifiers are effective for high-dimensional problems due to their modular structure. However, the conventional structure, where each module corresponds strictly to a single input feature, limits the model's explainability regarding the class-specific contributions and rule importance. This paper proposes a method to automatically reconstruct the module structure of multi-class SIRMs by clustering consequent weight vectors of pre-trained rules. We introduce two reconstruction approaches: the ``Indexed Method,'' grouping rules based on their class-specific contributions, and the ``Sorted Method,'' grouping rules based on their importance to classification. Experimental results on benchmark datasets demonstrate that the proposed approaches achieve accuracy on par with that of the conventional feature-based SIRMs. Furthermore, quantitative evaluations using the faithfulness and stability metrics indicate that the proposed methods enhance explainability and provide more reliable and robust explanations of the model's behavior.

GS23-2 A Recursive Reservoir Concatenation Algorithm with Median Absolute Deviation for Image Restoration

Takahiro Kawano and Michiharu Maeda
(Fukuoka Institute of Technology, Japan)

This paper describes an application of recursive reservoir concatenation (RRC) for removing random-valued impulse noise. RRC has been suggested for image restoration derived from reservoir computing for removing salt-and-pepper noise, in which some pixels in an image randomly change to black or white. Although the conventional approaches such as median filters, adaptive filters, and morphological processing show certain effectiveness. They often induce blurring of fine details and loss of edge information. In contrast, RRC utilizes the dynamic characteristics of reservoir computing and has the advantage of simultaneously achieving nonlinear mapping and noise removal. This study introduces noise position detection according to the median absolute deviation (MAD). MAD-based Detection is statistically robust indicator that can stably distinguish between normal and noisy pixels even under high noise density conditions. By introducing a threshold parameter in MAD, the sensitivity of noise detection is controlled, and the relationship between the threshold and the noise removal rate is systematically evaluated to examine the influence of detection accuracy and false detection rate on restoration performance. In numerical experiments, various noise densities are applied to standard benchmark images. Comparisons are made with median filtering, adaptive median filtering. The evaluation indices include peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM). Numerical results exhibit that RRC combined with MAD-based detection effectively removes random-valued impulse noise while preserving fine structures and edges.

GS23-3 Evaluating Region-Specific Generative AI Models for Reproducing Historical Townscape Features

Taichi Himegi, Nobuo Mishima, Mohamed Rami Derbel, Wen Liang Yeoh and Osamu Fukuda
(Saga University, Japan)

To preserve the landscapes of Important Traditional Buildings Preservation Districts, it is essential that building renovations and reconstructions harmonize with the surroundings. However, current design decisions rely heavily on the designer's subjective judgment, lacking objectivity. While recent advances in image-generating AI have enabled automated architectural design, generic models struggle to faithfully reproduce each district's unique styles, material textures, and atmosphere. This research aims to establish a method for training generative AI to learn regional characteristics, thereby generating architectural designs that accurately match historical landscapes, and to evaluate their fidelity both subjectively and objectively. Ten districts nationwide were selected as subjects, and a dataset was created by photographing building images and classifying them into six categories. Subsequently, using Stable Diffusion, we performed additional training with approximately 120 images per class and feature description prompts, constructing six region-specific models. Two evaluation methods were employed: human judgment of feature reproduction accuracy and objective assessment using an AI classification model (YOLOv8). The classification model was trained on data-augmented images to enhance its generalization performance. This paper quantitatively demonstrates, based on these evaluation results, the extent to which generative AI can reproduce regional landscape features. Furthermore, we examine the impact of prompt complexity on generation performance.

GS23-4 Enhancing Crop Yield Prediction Under Climate Change Using Hybrid Machine Learning Framework with Climatological Normalization

Kenshiro Kumagai, Ryo Hatano and Hiroyuki Nishiyama
(Tokyo University of Science, Japan)

As the effects of climate change intensify, interannual variability in crop yields has increased, making agricultural production more uncertain. Consequently, the demand for accurate yield prediction methods that can estimate production prior to harvest is growing, particularly to support food security policies and risk management strategies. Recent studies have demonstrated the potential of machine learning approaches; however, many remain limited to specific regions or crops, raising concerns regarding generalisability under diverse climatic conditions. In this study, we propose a crop yield prediction framework based on long short-term memory (LSTM) models for handling meteorological time-series data. The proposed framework incorporates an attention mechanism to emphasise critical growth periods, as well as a climatological-ratio-based gating strategy that explicitly accounts for the degree of meteorological anomalies. In addition, the effectiveness of a hybrid architecture integrating a gradient boosting machine (GBM) is systematically evaluated. The proposed framework is assessed across multiple crops and countries. Experimental results indicate that, while performance varies depending on crop and regional characteristics, the proposed approach outperforms baseline models in several cases. These findings demonstrate the effectiveness and partial generalisability of the framework for crop yield prediction under climate change conditions.

GS23-5 Disentangling Tonality and Function in Chord Progressions using Self-Supervised Learning

Toma Kawakita[1], Yasumasa Tamura[2] and Masahito Yamamoto[2]
([1]Graduate School of Information Science and Technology, Hokkaido University, Japan)
([2]Faculty of Information Science and Technology, Hokkaido University, Japan)

Chord progressions are a fundamental element of musical structure, yet computational modeling of chords faces a significant challenge: the entanglement of absolute pitch information (tonality) and relative harmonic roles (function). Standard representation learning approaches often fail to distinguish between physically identical chords that function differently depending on the context, or conversely, functionally equivalent chords that differ in pitch due to transposition. To address this, we propose a novel self-supervised learning framework that explicitly disentangles tonality and function within chord embeddings. Our method leverages the property that harmonic function is invariant to transposition, whereas tonality is determined by the specific pitch arrangement. We process original and transposed chord progressions in parallel using a Transformer-based encoder and optimize the model through a multi-task learning objective. This includes a proposed "Hybrid Embedding MLM" that swaps embedding components across transpositions, along with metric learning and adversarial training to enforce separation. Experimental results on the ChoCo dataset demonstrate that our model successfully acquires two distinct subspaces: a function embedding that captures relative roles (e.g., Roman numerals) independent of the key, and a tonality embedding that robustly tracks local keys and modulations. This disentangled representation offers superior interpretability and provides a robust foundation for music analysis tasks.

GS23-6 Influence of Basis Functions on Dictionary Learning with K-Singular Value Decomposition for Image Denoising

Ataru Shakagori and Michiharu Maeda
(Fukuoka Institute of Technology, Japan)

This paper examines the design of an initial dictionary for image denoising through dictionary learning with a foundation in sparse representation. Although dictionary learning has proven effective for image denoising, its performance is significantly influenced by the choice of the initial dictionary. Conventional approaches typically rely on dictionaries founded on the discrete cosine transform (DCT), and the selection of this initial dictionary plays a critical role in determining both the final denoising performance and the computational efficiency of the learning process. In this study, we investigate dictionary initialization with basis functions derived from the discrete sine transform (DST) and the discrete Hartley transform (DHT). We employ the K-singular value decomposition (K-SVD) algorithm for dictionary updates and compare the performance of our DST- and DHT-based initializations with the conventional DCT-based initialization. The objective of this work is to evaluate the extent to which the choice of initial dictionary affects learning in image denoising.

January 22 (Thursday), 09:00-10:45

B1 Meeting Room 4

GS28 Motion planning and navigation

Chair: Keigo Watanabe (IPU Tokyo, Japan / Okayama University, Japan)

GS28-1 Development of a mobile robot control system with static obstacle avoidance function using LiDAR sensor and depth camera

Tomoya Kihara, Yoshitaka Matsuda, Ruimin Wang, Takenao Sugi and Satoru Goto
(Saga University, Japan)

In this study, control system of a mobile robot with LiDAR sensor and depth camera for a static obstacle avoidance is developed. In the obstacle detection, the range of obstacle detection is extended by integrating data from both LiDAR sensor and depth camera. In the obstacle detection experiment using two static obstacles, both obstacles were detected at positions higher than those detectable by LiDAR sensor alone, and the system's ability to expand the obstacle detection range was confirmed. However, a problem was found since the depth camera's detection range was limited to objects at the same height as the camera itself, it could not detect obstacles positioned between the LiDAR sensors and the depth camera. Then, a system for avoiding static obstacles is constructed. The effectiveness of the developed system was confirmed by conducting static obstacle avoidance experiments. In the static obstacle avoidance experiment with 10 trials, the mobile robot stopped in front of the obstacle and failed at 3 trials. However, it successfully avoided the obstacle at 7 trials.

GS28-2 Autonomous mobile system using Adaptive resonance theory-based Topological Clustering for illuminance measurement robot

Tomoki Fujii[1], Koki Sakamoto[1], Makoto Tsujimoto[2], Kazuhiko Taniguchi[2], Yuichiro Toda[1] and Takayuki Matsuno[1]
([1]Okayama University, Japan)
([2]Kinden Corporation, Japan)

In recent years, due to the declining birthrate and aging population, labor saving through automation is expected in various fields and workplaces. Among these, measurement of illuminance at construction sites is one of the most desired tasks to be automated by robots, since it is a night-time and long-hour job. A method based on Growing Neural Gas with Different Topologies (GNG-DT), an unsupervised learning method that abstracts the environment into a graph structure suitable for path planning and enables learning of topological structures, has been proposed for automating illuminance measurement. However, GNG-DT has many parameters, and it is difficult to set the spacing between nodes arbitrarily due to the effect of node oscillation. In this study, we propose a map construction method based on Adaptive Resonance Theory-based Topological Clustering with Different Topologies, in which nodes are added based on distance thresholds, and the node spacing can be changed as needed. The proposed method is intended to improve the efficiency and stability of the navigation system in a semi-known environment by changing the node spacing as necessary. We demonstrate the effectiveness of the proposed method by conducting demonstration experiments using actual equipment in a semi-known environment.

GS28-3 Vision-Based Sim2Real Transfer of Map-Free Reinforcement Learning Navigation with Semantic Segmentation and Multi-Sensor Fusion

Daisuke Amano and Kazuyuki Morioka
(Meiji University, Japan)

This study proposes a vision-based Sim2Real transfer framework for map-free deep reinforcement learning (DRL) navigation, integrating semantic segmentation and multi-sensor fusion to achieve stable autonomous navigation without prior maps. The main contribution of this work is to demonstrate that a vision-based DRL policy trained entirely in simulation can be successfully transferred to real-world environments through an integrated perception and localization framework. The proposed system enables goal-directed navigation based on multi-sensor fusion of visual, inertial, and wheel-encoder information, providing a lightweight and scalable solution for low-cost mobile robots.

GS28-4 Sim-to-Real Mobile Robot Navigation with Action Models via Generative Adversarial Imitation Learning Based on Automatic Demonstration Data Recording

Kazuki Mishina and Kazuyuki Morioka
(Meiji University, Japan)

Imitation learning faces challenges in the efficiency of demonstration data collection and in surpassing expert performance. To address these issues, this study integrates automatic demonstration data recording with Generative Adversarial Imitation Learning (GAIL). Demonstration data are automatically collected in simulation environments using basic action models trained through curriculum learning, which serve as expert models for indoor navigation and eliminate the need for manual data collection. Using the recorded demonstration data, GAIL updates the action model. The training process encourages exploration beyond simple imitation. This prevents excessive convergence to expert trajectories and promotes the learning of more efficient navigation behaviors. The learned model is then applied to a ROS-based mobile robot, and its performance is evaluated using the goal success rate in both simulation and real environments to assess Sim-to-Real applicability. Experimental results demonstrate that the proposed method achieves higher performance than the expert model used for demonstration generation, while enabling stable long-distance navigation without path planning or predefined waypoints.

GS28-5 Particle Filter-Based Indoor Robot Localization with Global Position Estimation by Cross-Modal Representation of Floor Map and Monocular Camera Image

Yuto Hirakawa and Kazuyuki Morioka
(Meiji University, Japan)

Visual localization in indoor environments typically relies on pre-built 3D point cloud maps or heavy depth estimation models. However, these approaches often limit scalability and robustness against layout changes. To address these issues, this paper proposes a lightweight and scalable localization method that utilizes a 2D floor map and a monocular RGB camera without requiring depth sensors, LiDAR, or pre-scanning. We introduce a cross-modal representation learning framework based on contrastive learning, which embeds camera images and floor map patches into a shared latent space based on geometric structures. This approach avoids explicit edge extraction, making it robust to textureless surfaces. By integrating this learned observation model into a Monte Carlo Localization (MCL) framework, the proposed method effectively resolves perceptual aliasing through temporal filtering. Experimental results demonstrate that the proposed method achieves accurate localization in both seen and unseen environments, validating its generalization capability to novel layouts without retraining. This system offers a practical solution for autonomous mobile robots in changing indoor environments.

GS28-6 Fast, Smooth, and Collision-Free Path Generation with Multi-Stage Multi-Layered Perceptron and Modified STOMP

Minje Kim, Soma Fumoto and Takeshi Nishida
(The University of Kitakyushu, Japan)

This paper presents a hybrid motion-planning framework that integrates fast neural path inference, lightweight smoothing, and probabilistic refinement for articulated robots in high-mix, variable-volume production environments. The method first employs a Multi-stage Multilayer Perceptron (MMLP) to rapidly generate an initial sequence of intermediate waypoints. A Savitzky–Golay (SG) filter is then applied to enhance the geometric smoothness of the inferred path while preserving endpoint fidelity. Finally, a modified Stochastic Trajectory Optimization method, m-STOMP, performs local probabilistic adjustments near obstacles to ensure collision-free execution. Implemented in ROS 2 with a 6-DOF industrial manipulator, the proposed three-stage pipeline achieved a 100% success rate across 100 trials, with a low average computation time of 0.39 s. The approach produced shorter and smoother trajectories—measured in terms of total joint displacement—than RRT* and standalone MMLP. These results demonstrate the effectiveness of the method for efficient and reliable motion planning in complex three-dimensional robotic workspaces.

GS28-7 Improvement of Trajectory Optimization Algorithm for Path Reuse Method

Tsubasa Watanabe, Soma Fumoto and Takeshi Nishida
(The University of Kitakyushu, Japan)

We propose m-STOMP, a stochastic optimization method tailored for the Path Reuse (PR) method, to achieve fast and safe replanning in high-mix manufacturing. Conventional optimization methods, such as Stochastic Trajectory Optimization for Motion Planning (STOMP), often face a trade-off between computation time and stability. m-STOMP improves this in two ways. First, it starts the search closer to the solution by using a trajectory generated by the PR method that retains shape information, rather than a straight line. Second, it enhances computational efficiency and convergence stability by using automatic hyperparameter tuning for each joint. In simulations, m-STOMP demonstrated a higher planning success rate and reduced computation time compared to the baseline, while improving stability.

January 22 (Thursday), 13:00-14:15

3F Meeting Room 32

OS12 AROB: EcoMechatronics and AI

Organizers: Maki K. Habib (The American University in Cairo, Egypt)
Fusaomi Nagata (Sanyo-Onoda City University, Japan)

OS12-1 EcoMechatronics: Sustainable Design and Human-Centric Mechatronic Systems

Maki K. Habib[1] and Fusaomi Nagata[2]
([1]The American University in Cairo, Egypt)
([2]Sanyo-Onoda City University, Yamaguchi, Japan)

Conventional mechatronic systems have historically been optimized for performance, precision, and efficiency, often treating sustainability and human-centric considerations as secondary constraints. This design paradigm is increasingly misaligned with global sustainability imperatives and the human-centric vision of Industry 5.0. This paper introduces EcoMechatronics as a structured design framework that embeds ecological responsibility and human-centricity directly into mechatronic system architecture, control, and lifecycle management. Using a comparative conceptual synthesis across robotics, manufacturing, and energy-aware automation, three guiding principles are identified: lifecycle-oriented design, resourceefficient integration, and human-centered adaptability. The paper clarifies how these principles extend traditional mechatronics by transforming sustainability and human factors from post-design evaluations into primary system objectives. Illustrative applications in robotic disassembly, hybrid manufacturing, and energy-aware automation demonstrate the domain-agnostic relevance of the framework. The paper concludes by discussing adoption barriers and research gaps, emphasizing the need for sustainability-aware benchmarks, lifecycle verification, and governance alignment to support the transition toward Industry 5.0.

OS12-2 Learning-less Dataset Model for Leaf Disease Detection and Similar Symptom Search

Yuji Peng[1], Kiri Yamaguchi[2], Fusaomi Nagata[2], Zhelin Zheng[2], Keigo Watanabe[3] and Maki K. Habib[4]
([1]Graduate School of Engineering, Sanyo-Onoda City University, Japan)
([2]Sanyo-Onoda City University, Japan)
([3]Okayama University, Japan)
([4]The American University in Cairo, Egypt)

Early detection of plant diseases is crucial for preventing yield loss and ensuring crop quality. This study introduces a learning-less dataset model for leaf disease detection and symptom-based image retrieval. The method leverages pre-trained CNN features and evaluates similarity using Kullback Leibler divergence (KLD), cosine distance, and L2 norm, eliminating the need for initial model training from scratch. Experiments on tomato leaf disease datasets demonstrate that the proposed approach achieves competitive accuracy while substantially reducing computational cost compared to conventional CNN-based methods.

OS12-3 Systematic Evaluation of Variational Autoencoder Training Process with Independent Control of Encoder and Decoder

Zhelin Zheng[1], Fusaomi Nagata[1], Akimasa Otsuka[1], Hirohisa Kato[1], Keigo Watanabe[2], Maki K. Habib[3] and Ahmad Shahrizan Abdul Ghani[4]
([1]Sanyo-Onoda City University, China)
([2]Okayama University, Japan)
([3]The American University in Cairo, Egypt)
([4]University Malaysia Pahang Al-Sultan Abdullah, Malaysia)

In Variational Autoencoders (VAEs) training, balancing the Kullback Leibler divergence (KLD) loss and reconstruction loss calculated by mean squared error (MSE) loss is crucial for achieving stable convergence and effective latent representation learning. However, the simultaneous optimization of these two losses often leads to unstable training or redundant computation time due to conflicting gradients. To address this issue, this paper investigates the influence of the repeat number, defined as the number of consecutive optimization steps for each loss, on the convergence behavior and performance of VAEs.

OS12-4 Development of Brain Organoid Stimulation Device for Quadruped Robot Using Spiking Central Pattern Generator Circuit

Shumpei Nagano[1], Shuxin Lyu[1], Yoshiho Ikeuchi[2], Yoshio Mita[3] and Ken Saito[4]
([1]Department of Precision Machinery Engineering, Graduate School of Science and Technology, Nihon University, Japan)
([2]Institute of Industrial Science, The University of Tokyo, Japan)
([3]Dept. of Electrical Engineering and Information Systems (EEIS), School of Engineering, The University of Tokyo, Japan)
([4]Department of Precision Machinery Engineering, College of Science and Technology, Nihon University, Japan)

To understand the mechanism of gait generation in animals, the authors are researching robot control using Hardware Spiking Neural Networks (HSNNs), analog circuits that mimic biological nervous systems. To achieve more biologically plausible control, we are also exploring the use of brain organoids derived from human induced pluripotent stem cells. Delivering gait information to brain organoids requires external stimulation; therefore, a signal-generating circuit is essential. To address this, we developed a brain organoid stimulation device using a spiking Central Pattern Generator (CPG) composed of HSNNs. This paper describes the device developed, which integrates a CPG circuit, a peripheral circuit, and a microcontroller. Measurement results confirm that the CPG circuit generates quadrupedal gait patterns triggered by external single-pulse waveform inputs. Furthermore, the circuit operates at frequencies to which brain organoids may respond, such as 1 Hz. These results demonstrate that the developed device can deliver stimulation signals appropriate for brain organoids.

OS12-5 Design Support System for Real-Time Object Detection Model and Instance Segmentation Model

Naoki Nishida[1], Fusaomi Nagata[1], Kaito Otsu[1], Akio Kurozumi[1], Keigo Watanabe[2] and Maki K. Habib[3]
([1]Sanyo-Onoda City University, Japan)
([2]Okayama University, Japan)
([3]The American University in Cairo, Egypt)

This article reports on newly implemented functionalities that enable students to experience the design of YOLO model for real-time object detection and as well as the SOLO and Mask R-CNN models for instance segmentation, which are expected to be applied to autonomous driving systems, automated manufacturing lines, medical diagnostic systems, and even mobile robots in everyday life. It is demonstrated trough experiments that students can easily learn how to train and test YOLOX models, SOLOv2 models, and Mask-R-CNN models using the proposed system.

January 22 (Thursday), 13:00-14:30

3F Meeting Room 33

OS13 AROB: Intuitive Human-System Interaction I

Organizer: Masao Yokota (Fukuoka Institute of Technology, Japan)

OS13-1 Cognitive Function Support System using Gaze Control

Takato Ikeda and Tomoyuki Ishida
(Fukuoka Institute of Technology, Japan)

The number of people with dementia in Japan is increasing, creating various social problems, such as the burden of care and financial burden on patients and their families. Consequently, systems capable of continuously assessing and maintaining cognitive function in daily life are becoming increasingly important. Recently, virtual reality (VR) technology has enabled natural simulation of cognitive functions, such as memory, attention, and judgment, in highly immersive environments. Therefore, this study developed a VR-based cognitive function support system that incorporates eye-tracking technology to assess cognitive function easily and objectively. The proposed system combines an immersive VR experience with high-precision gaze analysis to evaluate cognitive function and quantitatively understand changes resulting from training. This approach provides continuous feedback and improvement based on each user fs cognitive state. By establishing an autonomous cognitive support model that integrates assessment and prevention, this study aims to contribute to a society in which people of all ages can independently manage their cognitive function daily.

OS13-2 Laser Harp System using Virtual Reality Technology

Harui Cho and Tomoyuki Ishida
(Fukuoka Institute of Technology, Japan)

In recent years, with the wider availability and improved performance of head-mounted displays, virtual reality (VR) technology has emerged as a promising platform for new musical experiences. In this study, we developed a VR laser harp system specialized for playing melody lines, aimed at reducing the technical barriers inherent in traditional instrument learning and considerably improving learning efficiency in music education. Traditional instrument learning requires substantial time and resources to master performance techniques, which can lower motivation. Physical laser harps offer an intuitive interface but require delicate hand positioning, making precise melodic performance difficult. To address these issues, we created a VR laser harp environment with an intuitive performance interface. This study demonstrates the educational potential of VR instruments and contributes to enhancing music learning efficiency by providing a practical and effective solution for melody acquisition, a core component of music learning. Furthermore, this study offers new insights into the importance of practical functionality in digital instruments for educational applications, contributing to the integrated development of STEM (Science, Technology, Engineering, and Mathematics) education and music learning.

OS13-3 Development and Evaluation of a Facilitation Robot for Supporting Idea Generation — The Future of Ideation through Creative Support Robotics —

Tetsuya Shiba[1], Hideki Muto[1], Tetsuro Ogi[1], Yuki Kida[1], Linqi Jin[1], Sota Saito[1] and Takashi Maeno[2]
([1]Keio University, Japan)
([2]Musashino University, Japan)

This study aims to support creative idea generation by developing and evaluating a facilitation robot that integrates artificial intelligence (AI) and robotics. While the importance of creative discussion has increased, conventional facilitation practices heavily depend on the experience and personal skills of expert facilitators, making reproducibility a major challenge. To address this issue, we designed a robot that supports idea generation by sharing ideation methods, providing dynamic questioning, and encouraging participant participation, and conducted a comparative experiment with a human facilitator. The experiment employed a counterbalanced design in which the order of facilitation conditions and task themes was reversed. The results showed that, in the initial session, the human facilitator was superior in some aspects of the ideation process, whereas in repeated sessions the differences between conditions diminished, and the robot condition achieved a level of ideation experience comparable to that of the human facilitator. These findings suggest the feasibility of dialogue-driven, reproducible facilitation and highlight a new direction for creativity support through human–robot collaboration.

OS13-4 Evaluation on sense of presence in Online VR Communication by spatial audio

Akito Nishi and Yasuo Ebara
(Osaka Electro-Communication University, Japan)

Compared to commonly used online communication methods, online VR communication enables communication by a high sense of presence in a virtual space with its high-quality visual representation afforded by the wide field of view of VR headsets. However, in communication using VR headsets, the decline in sense of presence due to discrepancies between visual and auditory information has been a concern. To improve the decline in sense of presence, we studied to realize tele-immersion technology using VR. In existing studies, we developed an online VR communication system that applies HRTFs in real-time to participants' conversation voices according to the coordinate relationship and the relationship of face orientation between avatars in the virtual space, allowing the voice of the conversation partner in the virtual space to be heard spatially. In addition, we evaluated the sense of presence achieved by implementing spatial sound in online VR communication, as well as the feelings of pleasantness/unpleasantness and changes in personal space during proximal interactions. In this study, we conduct a detailed evaluation of the effect of spatial audio on the sense of presence in online VR communication, which could not be clearly assessed using previous evaluation metrics. This evaluation is based on psychological data from subjective questionnaire surveys and physiological data from time-series changes in heart rate during the experiment.

OS13-5 Implementation of motorcycle recommendation system for beginners using user review data

Yosuke Minato and Yasuo Ebara
(Osaka Electro-Communication University, Japan)

The number of domestic motorcycles demand has been declining since its peak in 1982, attributed to factors such as stricter regulations, economic recession, and inadequate infrastructure for motorcycles. On the other hand, recovering domestic motorcycles demand requires attracting new users, it is important to resolve the challenges for core beginners. However, it is difficult for beginners who lack knowledge of motorcycle specifications and performance to choose the optimal motorcycle for them. In this study, we implemented a motorcycle recommendation system that uses review data to be easy to use even for beginners. This system extracts information from a database based on conditions entered by the user and presents the most suitable motorcycles in a ranked format. Furthermore, we quantitatively evaluated the effectiveness of this system through a questionnaire evaluation. From the evaluation results, this system was able to present, on average, a top-two most suitable motorcycle, and it was confirmed that the operability was also highly evaluated.

OS13-6 Implementation and Evaluation of an LED Based Emotional Expression in IoT Avatars

Toshihiko Chiba, Yuki Kida and Tetsuro Ogi
(Graduate School of System Design and Management, Keio University, Japan)

IoT avatars, which transform everyday objects into remotely operated communication media, have recently been proposed as a alternative to conventional avatar robots. While such approaches reduce deployment constraints related to cost and space, ordinary objects lack facial expressions and bodily gestures, making emotional communication difficult. This limitation poses a challenge for achieving smooth and intuitive interaction comparable to human-to-human communication. To address this issue, we developed an IoT avatar system that enables seamless transmission of a remote operator's emotional state to a physical object. The system detects the operator's joy using facial expression analysis and speech-based emotion recognition and expresses the detected emotion through LED-based visual feedback embedded in the object. Specifically, LEDs mounted on a disinfectant bottle avatar were modulated with heartbeat-like blinking patterns to convey emotional states. An evaluation experiment investigated how LED color and emotion detection modality influenced users' perception of the avatar's emotional expression. The results indicate that emotional expression through LEDs is feasible, but its effectiveness strongly depends on color selection. These findings suggest that designed LED-based expressions may contribute to intuitive emotional communication in IoT avatars, even in the absence of humanoid features.

January 22 (Thursday), 13:00-14:30

B1 Meeting Room 1

OS15 AROB: Robotics with Intelligence and/or Informatics I

Organizers: Tetsuya Kinugasa (Kindai University, Japan)
Masatoshi Hatano (Nihon University, Japan)

OS15-1 Determination of an Optimal Posture to Improve Hand Generating Forces Using Genetic Algorithm on Two-Dimensional Irregular Terrain

Shimon Matsuzaki[1], Tomoki Hirano[2], Keitaro Kawai[2] and Masatoshi Hatano[3]
([1]Nihon University, Graduate school, Japan)
([2]Nihon University, Under Graduate school, Japan)
([3]Nihon University, Japan)

The purpose of this research is to propose a determination method of an optimal posture of a rescue robot for improving the hand force generated by driving torques of sub-crawlers. In collapsed buildings caused by disasters, scattered debris may prevent the door from opening. At this time, hand generating forces by only the mounted manipulator of the rescue robot are often insufficient. On the other hand, the rescue robot has sub-crawlers to change postures of the robot body. Then, we propose a method to utilize torques generated by sub-crawlers to add forces of the hand through link mechanisms for opening broken doors. We also propose a method to determine the optimal posture of a rescue robot using a genetic algorithm (GA) in order to solve the problem of the redundancy. It is shown that our proposed method can derive an optimal posture on two-dimensional irregular terrain with the GA. In addition, the validity of the method was clarified with the actual experiments.

OS15-2 Developments of Sampling Mechanisms for Sediment Collection Robots in Tidal Flats

Riki Sugama[1], Shimon Matsuzaki[2], Masaki Yamazaki[1], Tomoki Hirano[2] and Masatoshi Hatano[3]
([1]Nihon University,Undergraduate school, Japan)
([2]Nihon University, Graduate school, Japan)
([3]Nihon University, Japan)

The purpose of this research is to develop a robotic system capable of collecting bottom sediment while traveling on muddy tidal flats. Tidal flats and sandy beaches play important roles in disaster prevention and ecosystem support, yet 13% of sandy beach area was reportedly lost in the 15 years up to 1998, and erosion continues worldwide. This erosion is caused by sand drifting offshore, and unless the lost sand is compensated for, beaches will keep shrinking. Factors include climate-change-driven current shifts, coastal development, and reduced river sediment due to dams, though the precise mechanisms remain unclear. To model sand drift mathematically, it is first necessary to identify sediment characteristics in target coastal areas. Currently, researchers must walk dangerously on tidal flats to collect samples, and only a few points can be sampled. Therefore, robotic sampling is required. We developed a robot and a new sediment sampling mechanism, and this study investigates how edge-shape differences—such as claw geometry—affect the amount of mud collected. Experimental results are presented and discussed.

OS15-3 Research on Harvesting Robot Systems with Real-Time Grading Functions and Visual Servo Controls

Junha Hwang[1], Xi Liu[1], Keito Arai[2], Kimiyo Matsukura[3], Ken Saito[3], Minami Kaneko[3], Kai Masuda[3], Shinyou Muto[3], Wataru Imahayashi[3], Syota Katagiri[3] and Masatoshi Hatano[3]
([1]Nihon University, Under Graduate school, Japan)
([2]Nihon University, Graduated school, Japan)
([3]Nihon University, Japan)

This study aims to develop a robotic harvesting system that automates agricultural picking tasks by estimating the position, size, and grade of crops using AI-based image analysis and controlling the robot through visual servoing. In Japan, labor shortages and the aging agricultural workforce have intensified, creating an urgent need to reduce the physical burden of farm work. Although smart farming initiatives are being promoted nationally, fully developed robots for fruit vegetables such as eggplants and tomatoes remain limited. Unlike root vegetables, which can be harvested collectively, fruit vegetables require selective harvesting due to variations in maturity and spatial distribution. This research focuses on eggplants, a crop with high annual production and slow automation progress. Since about 75% of Japanese farms are small family-run operations, reducing labor must include not only harvesting but also post-harvest sorting. Therefore, this study develops a robotic system capable of harvesting eggplants while considering their size and quality grade.

OS15-4 Development of a Camera Assistant Robot for Pet Surgery

Yuki Sukeyasu[1], Kensho Yamagata[1] , Shimon Matsuzaki[2], Junha Hwang[1], Kazuya Edamura[3], Hiroharu Yamashita[3] and Masatoshi Hatano[3]
([1]Nihon University, Under Graduate school, Japan)
([2]Nihon University, Graduated school, Japan)
([3]Nihon University, Japan)

The purpose of this study is to develop a compact camera assistant robot for pet surgery. Existing surgical assist robots (da Vinci, hinotori) are highly expensive, making them difficult to implement in small or medium-sized medical institutions. To address this, we focus on developing a low-cost robot specialized in camera assistance, enhancing practicality without attempting to replace all human roles in the operating room. The proposed robot is equipped with artificial intelligence (AI), enabling it to learn and recognize surgical target areas and autonomously track them during procedures. This functionality eliminates the need for manual camera adjustments, ensuring a stable surgical field. In veterinary surgery, where operating space is limited, the compact design of the robot is critical. Its small form factor minimizes interference with the surgeon's hand movements and instruments, contributing to enhanced surgical safety. Furthermore, the robot moves the camera along a conical trajectory centered around a small incision, allowing wide field-of-view coverage while reducing invasiveness. A prototype of the AI-powered camera assistant robot was developed, and its basic motion capabilities were verified through experiments.

OS15-5 CentiPLAY: A Super-Multi-Legged Robot Realizing 3D Leg Motion and Ground Reaction Force Perception via Mechanical Constraints

Tomoya Nishiyama[1], Tatsuki Honjo[1], Manato Shiotani[2], Ryota Hayashi[2], Koji Yoshida[2] and Tetsuya Kinugasa[3]
([1]Department of System Science, Graduate School of Science and Engineering, Okayama University of Science, Japan)
([2]Department of Mechanical Systems Engineering, Faculty of Engineering, Okayama University of Science, Japan)
([3]Department of Robotics, Faculty of Engineering, Kindai University, Japan)

Myriapods such as centipedes and millipedes exhibit adaptive locomotion through coordinated interactions among their numerous body segments and legs. Understanding these mechanisms is essential for advancing the development of high-performance multi-legged robots. In this study, we developed CentiPLAY-P, a compact myriapod-inspired robot capable of generating three-dimensional leg motions and perceiving ground reaction forces through simple mechanical structures. Each leg incorporates one passive roll joint and one active yaw joint driven by a single servo motor. A curved slide rail combined with a ratchet mechanism converts the periodic yaw oscillation into continuous three-dimensional leg trajectories. An elastic element and a potentiometer installed on the roll axis enable the robot to detect ground reaction forces and distinguish differences in terrain. By varying the oscillation frequency, phase differences among legs, and duty factors, the robot successfully reproduces various gait patterns observed in real myriapods. Experiments conducted on different surfaces, such as carpet and desktop environments, revealed distinct roll-angle responses, indicating environment-dependent sensory reactions. These results demonstrate that CentiPLAY-P serves as a practical platform for investigating myriapod locomotion and contributes to the development of multi-legged robots with adaptive mobility.

OS15-6 Empirical Analysis of Swing-Leg Impact Dynamics in Relation to Stability in Passive Dynamic Walking

Kazuto Matsuura[1], Ryota Hayashi[2], Koji Yoshida[2], Tetsuya Kinugasa[3] and Yasushi Iwatani[3]
([1]Department of System Science, Graduate School of Science and Engineering, Okayama University of Science, Japan)
([2]Department of Mechanical Systems Engineering, Faculty of Engineering, Okayama University of Science, Japan)
([3]Department of Robotics, Faculty of Engineering, Kindai University, Japan)

Passive dynamic walking is a phenomenon in which a walking machine without actuation descends a slope by exploiting its mechanical properties and interactions with the environment. Previous studies have shown that walking stability is influenced by factors such as mass distribution, foot geometry, and collision behavior at swing-leg contact. However, many studies have focused on walkers with relatively large circular-arc feet, and the stability characteristics of walkers with small arc radii—where instability and falling are more likely—have not been sufficiently verified through experiments using physical machines. In this study, a passive dynamic walker equipped with small circular-arc feet was developed, and experimental analyses were conducted to investigate the effects of hip mass, leg center-of-mass position, and swing-leg collision tendencies on walking stability. The results demonstrated that increasing hip mass improved stability, and that stable walking was characterized by alternating moderate forward and backward collisions. In contrast, unstable walking tended to involve forward collisions during swing-leg contact, accompanied by excessive backward impacts immediately before falling. While previous research suggested that backward collisions contribute to stability, the present findings indicate that when the foot-arc radius is small, pronounced forward-collision tendencies are closely associated with reduced walking stability.

January 22 (Thursday), 13:00-14:15

B1 Meeting Room 2

OS17 AROB: Swarm and Modular Robotics I

Organizer: Ryusuke Fujisawa (The University of Kitakyushu, Japan)

OS17-1 Aerial Interception of Multiple Drones by Multiple Distributed Chasers: Effect of spreading factors

Sutthiphong Srigrarom[1], Nontaphat Charuvajana[1], Panithan Rithburi[2], Boo Cheong Khoo[1] and Florian Holzapfel[3]
([1]National University of Singapore, Singapore)
([2]Technical University of Munich Asia, Singapore)
([3]Technical University of Munich, Germany)

In this paper, we investigate the problem of air-to-air drone interception forcounter-UAS operations, using strategic surrounding approaches. The intruding drones may come in large numbers and are modeled by Boids Algorithm, analogous to flocks of birds. They assumed to destine towards the goal, of which the interceptors need to protect. The fleet of interceptors (agents) are assumed to be limited numbers and are equal to or smaller than the intruding drones. At the start, we employed a greedy strategy as a task-allocation method to break up, herd, and encircle the intruders. This framework is modeled as matching and optimization problem. The preliminary mix-integer, non-linear problem (MINLP) formulations are based on probability of interception and resource readiness. Once the interceptors conceptually notice the presence of the interceptors within their sensing ranges, the intruders could exhibit spreading behavior, which result in disintegrating the flock or chaotic evasive motions, and that, their grouping change. As a result, the interceptors need to redistribute their task assignments, and perhaps, be deployed in sequences of batches (groups) such that they can cover larger area, and allow more aggressive follow-up action. Preliminary results show that when the intruders start spreading, the motion of the interceptors are more chaotic and take longer time to intercept.

OS17-2 Synchronization-Enhanced RGB-D SLAM and Navigation for Quadruped Robots Using Intel RealSense D435i and RTAB-Map Under ROS 2

Weihua Li, Xiaoya Ni, Wendong Zhang and Sutthiphong Srigrarom
(National University of Singapore, Singapore)

This work presents the design and implementation of a complete RGB-D SLAM mapping and navigation system for a quadruped robot using an Intel RealSense D435i and RTAB-Map under ROS 2 Humble on Ubuntu 22.04. The proposed framework enables real-time 2D/3D environmental mapping, localization, and global path planning on the Unitree Go2 platform by integrating synchronized depth, RGB, and odometry data streams together with an A*-based navigation module. A common challenge in RGB-D SLAM for legged robots—the timestamp mismatch between the depth camera and onboard odometry, which frequently causes map distortion or localization drift during dynamic motion—is addressed through a custom synchronization mechanism that aligns the frame rates and temporal origins of the RealSense and Unitree odometry.

OS17-3 Research foundation and roadmap for autonomous transition AI and reinforcement learning for modular robots

Tomoya Negoro[1], Ryota Kinjo[1], Kenichiro Satonaka[1], Ryuusei Nishii[1], Ohashi Seiichi[1], Yuki Takagi[1], Hiroshi Oku[1], Yuki Tanigaki[1], Tomohiro Shimomura[1], Guang Yang[1], Xixun Wang[1], Fumitoshi Matsuno[1], Koki Harada[2] and Ryo Ariizumi[3]
([1]Osaka Institute of Technology, Japan)
([2]Nagoya University, Japan)
([3]Tokyo University of Agriculture and Technology, Japan)

The Moonshot project, "Self-Regenerative AI Robots for Lunar Exploration and Base Construction," aims to realize reconfigurable modular robots for uncharted environments. However, the "combinatorial explosion" of potential states and transition paths makes searching for optimal routes extremely difficult. To address this, we developed Canonical Robot Adaptive Graph Encode (CRAGE), which converts structural graphs into unique, normalized strings by leveraging geometric symmetry. CRAGE enables rapid isomorphism detection through string comparison, facilitating the generation of large-scale datasets for high-precision Graph Neural Networks (GNNs). Additionally, we established a geometric validation system that converts CRAGE data into URDF via GGDL, utilizing ROS 2 and MoveIt to eliminate physically infeasible configurations caused by collisions or joint constraints. Furthermore, we are implementing Reinforcement Learning (RL) to optimize transition stability and torque efficiency in physical simulations. By feeding RL-generated data back into GNNs to improve physical feasibility prediction, this research integrates CRAGE, GNNs, and geometric validation into a comprehensive AI framework for the autonomous transition of modular robots.

OS17-4 Continuous Transition Control Architecture for Self-Reconfiguring Modular Robots

Seiichi Ohashi, Ryota Kinjo, Kenichiro Satonaka, Ryuusei Nishii, Tomoya Negoro, Guang Yang, Xixun Wang and Fumitoshi Matsuno
(Osaka Institute of Technology, Japan)

Self-reconfiguring modular robots require autonomous reconfiguration capabilities, yet existing robotics frameworks assume static configurations. Since our robotic operation platform uses ROS 2, we initially developed a MoveIt-based geometric validation system that evaluates module connection feasibility using inverse kinematics, collaborating with CRAGE for combinatorial transition pathway generation and GGDL for URDF conversion. However, MoveIt's immutable RobotModel prevents runtime structure updates, forcing full reinitialization after each transformation. To overcome this limitation, we propose a novel architecture that converts URDF hierarchical structures into a flattened link and joint collection independent of the original hierarchy, enabling on-demand kinematic chain reconstruction. Implemented in Rust with Bevy as the 3D framework and Rapier for geometric validation, our system employs a Jacobian-based iterative solver for inverse kinematics computation. Performance evaluation demonstrates 16× faster initialization (0.6s vs. 9.6s) and 3.7× lower memory consumption compared to MoveIt. Our approach eliminates reinitialization requirements between transformation steps; for a 10-step transition, this reduces cumulative initialization time from 96 seconds (MoveIt: 9.6s × 10 steps) to 0.6 seconds (initial startup only), eliminating 95.4 seconds of overhead. The validated transition pathways provide training data for AI-driven autonomous reconfiguration. Future work will focus on control automation and hardware validation for lunar exploration applications.

OS17-5 Development of a GNN platform for transforming modular robots

Ryusei Nishii[1], Ryota Kinjo[1], Kenichiro Satonaka[1], Seiich Ohashi[1], Tomoya Negoro[1], Hiroshi Oku[1], Yuki Tanigaki[1], Guang Yang[1], Xixun Wang[1], Ryo Ariizumi[2] and Fumitoshi Matsuno[1]
([1]Osaka Institute of technology, Japan)
([2]Tokyo University of Agriculture and Technology, Japan)

This study addresses autonomous transition path planning for modular robots designed for lunar exploration. Identifying optimal paths is difficult because the search space becomes enormous as the number of modules increases. We propose a GNN framework to evaluate robot structures represented as graphs. We constructed two models: a "Transformability Judgment Model" to assess topological connectivity and a "Geometric feasibility judgment Model" to verify physical buildability. While the geometric model initially faced challenges with data bias, incorporating explicit negative examples significantly improved performance. Integrating these models allowed for transformability judgments based on geometric consistency, achieving approximately 98% accuracy. Specifically, the results confirmed that the GNN can judge a transformation as untransformable when it involves a geometrically infeasible structure, even if it is transformable in terms of the graph topology. This integration effectively eliminates geometrically infeasible paths from the search space. This approach efficiently narrows the search space for feasible autonomous transitions.

January 22 (Thursday), 13:00-14:30

B1 Meeting Room 3

OS19 AROB: Swarm-Machine Interaction: Models and Foundations for Collective Behavior I

Organizers: Hiroaki Kawashima (University of Hyogo, Japan)
Takayuki Niizato (University of Tsukuba, Japan)
Hitoshi Habe (Kindai University, Japan)

OS19-1 Tracking Feral Horses in Aerial Video Using Oriented Bounding Boxes

Saeko Takizawa[1], Tamao Maeda[2], Shinya Yamamoto[3] and Hiroaki Kawashima[1]
([1]University of Hyogo, Japan)
([2]The Graduate University of Advanced Studies (SOKENDAI), Japan)
([3]Kyoto University, Japan)

The social structures of group-living animals such as feral horses are complex and remain insufficiently understood, even within a single species. To investigate group dynamics, aerial videos are often utilized to track individuals and analyze their movement trajectories, which are essential for evaluating inter-individual interactions and comparing social behaviors. Accurate individual tracking is therefore crucial. In multi-animal tracking, axis-aligned bounding boxes (bboxes) are widely used; however, for aerial top-view footage of entire groups, their performance degrades due to complex backgrounds, small target sizes, high animal density, and varying body orientations. To address this issue, we employ oriented bounding boxes (OBBs), which include rotation angles and reduce unnecessary background. Nevertheless, current OBB detectors such as YOLO-OBB restrict angles within a 180° range, making it impossible to distinguish head from tail and often causing sudden 180° flips across frames, which severely disrupts continuous tracking. To overcome this limitation, we propose a head-orientation estimation method that crops OBB-centered patches, applies three detectors (head, tail, and head-tail), and determines the final label through IoU-based majority voting with confidence-based tie breaking. Experiments using 299 test images show that our method achieves 99.3% accuracy, outperforming individual models, demonstrating its effectiveness for robust OBB-based tracking.

OS19-2 Detection and Identification of Penguins Using Appearance and Motion Features

Kasumi Seko, Hiroki Kinoshita, Raj Rajeshwar Malinda and Hiroaki Kawashima
(University of Hyogo, Japan)

In animal facilities, continuous surveillance of penguins is essential yet technically challenging due to their homogeneous visual characteristics, rapid and frequent posture changes, and substantial environmental noise such as water reflections. In this study, we propose a framework that enhances both detection and identification performance by integrating appearance and motion features. For detection, we adapted YOLO11 to process consecutive frames to overcome the lack of temporal consistency in single-frame detectors. This approach leverages motion cues to detect targets even when distinct visual features are obscured. Our evaluation shows that fine-tuning the model with two-frames inputs improves mAP@0.5 from 0.922 to 0.933, outperforming the baseline, and successfully recovers individuals that are indistinguishable in static images. For identification, we introduce a tracklet-based contrastive learning approach applied after tracking. Through qualitative visualization, we demonstrate that the method produces coherent feature embeddings, bringing samples from the same individual closer in the feature space, suggesting the potential for mitigating ID switching.

OS19-3 YOLO Based Body Axis Estimation for Individual Fish with Temporal History Based Pose Correction

Shoma Kamata[1], Kohei Ohashi[1], Funo Suzuki[1], Hikaru Yamano[1],Taiki Nishimura[1], Hiroaki Kawashima[2], Hitoshi Habe[3] and Takayuki Niizato[1]
([1]Institute of Systems and Information Engineering, University of Tsukuba, Japan)
([2]Graduate School of Information Science, University of Hyogo, Japan)
([3]Faculty of Informatics, Kindai University, Japan)

This study proposes a correction method aimed at improving the accuracy of individual pose estimation in the analysis of schooling behavior in fish. First, we simultaneously estimate the poses of multiple individuals using YOLO v11, and apply anomaly detection based on the Local Outlier Factor (LOF) and the statistical distributions of inter-joint angles and segment lengths to identify erroneously detected keypoints. For keypoints judged to be abnormal, we perform corrections using a Kalman filter, the Rauch–Tung–Striebel (RTS) smoother, and a graph neural network (GNN), thereby improving pose estimation accuracy while preserving structural consistency. The overall system performance is evaluated using the mean squared error (MSE) with respect to ground-truth data, confirming the effectiveness of the proposed correction method. Furthermore, we discuss as future work the possibility of leveraging the characteristic periodic motions of fish in time-series data for additional corrections. The proposed method is expected to enable highly accurate pose estimation of individuals within a school—previously a challenging task—and to contribute to precise analysis of inter-individual interactions.

OS19-4 Towards Robust Multi-Object Tracking in Animal Groups with Particle Filters

Seiji Ishida[1], Issei Sasaki[2], Emyo Fujioka[2], Hiroaki Kawashima[3], Shizuko Hiryu[2] and Hitoshi Habe[1]
([1]Kindai University, Japan)
([2]Doshisha University, Japan)
([3]University of Hyogo, Japan)

Tracking animal groups remains a highly challenging problem due to strong appearance similarity among individuals, frequent occlusions, and irregular motion patterns that violate the linear motion assumptions underlying conventional Kalman filter–based methods such as ByteTrack and OC-SORT. To address these challenges, we propose a robust tracking framework that integrates high-precision object detection using YOLO with an adaptive particle filter capable of probabilistically modeling complex and non-linear motion. Unlike linear motion models, the proposed adaptive particle filter dynamically adjusts the particle distribution according to the target's motion velocity, enabling effective tracking of abrupt maneuvers such as sharp turns and rapid accelerations. Experimental evaluations conducted on nighttime thermal footage of bats and overhead videos of fish schools, using the MOTA and IDF1 metrics, demonstrate that the proposed method achieves superior tracking accuracy and stable identity preservation compared to existing multi-object tracking approaches, even under severe occlusion conditions. These results indicate that the proposed framework has significant potential for applications in ecological analysis and behavioral science.

OS19-5 Enhancing 3D Reconstruction Accuracy of Bat Flight Trajectories with Graph Matching

Ryota Uenaka[1], Issei Sasaki[2], Emyo Fujioka[2], Hiroaki Kawashima[3], Shizuko Hiryu[2] and Hitoshi Habe[1]
([1]Kindai University, Japan)
([2]Doshisya University, Japan)
([3]University of Hyogo, Japan)

Bats rely on echolocation to navigate and hunt, making accurate three-dimensional reconstruction of flight trajectories essential for behavioral analysis. Video-based stereo tracking has emerged as a non-invasive alternative to microphone arrays, and Kohyama et al. proposed a stereo-camera–based 3D tracking method. However, in practical recordings, 2D trajectories frequently become fragmented due to occlusion, scene clutter, and limited image resolution, and heavy observation noise further destabilizes viewpoint correspondence. These factors often lead to incorrect stereo matching and degrade the quality of 3D reconstruction. To address these challenges, this study introduces a graph-based matching framework that integrates GNN-based similarity estimation and temporal–geometric consistency constraints, enabling robust many-to-many association of fragmented 2D tracklets. Unlike traditional epipolar-only matching, the proposed method jointly exploits multiple motion and geometric cues through graph neural message passing, making correspondence more resilient to fragmentation and noise. To evaluate its effectiveness, we conducted extensive simulation experiments using synthetic bat flight data that include large positional noise, outliers, ghost tracks, and severe fragmentation patterns resembling real bat recordings. The proposed method consistently improved 3D reconstruction accuracy—particularly in MOTA, IDF1, and Recall—compared with a baseline stereo-matching approach, demonstrating its robustness under challenging tracking conditions.

OS19-6 Bidirectional Matching Method for Multi-Object Tracking of Individuals in Small Fish Schools

Kyoichiro Kawamura[1], Takayuki Niizato[2], Hiroaki Kawashima[3] and Hitoshi Habe[4]
([1]Graduate School of Kindai University, Japan)
([2]University of Tsukuba, Japan)
([3]University of Hyogo, Japan)
([4]Kindai University, Japan)

In recent years, Multi-Object Tracking (MOT) in video has become important in many fields. However, in small fish schools filmed from a top-down view, maintaining ID consistency is difficult for conventional online trackers using only Bounding Boxes (BBox) due to the elongated shape of individuals and frequent dense aggregation. In this study, to achieve high ID consistency while maintaining online capability, we propose a semi-online Bidirectional Matching method that utilizes past and future frames within a short time window. In the proposed method, we first perform tentative BBox tracking using OC-SORT, and subsequently re-evaluate trajectories and re-assign IDs within the window based on pose features derived from keypoints of the head, trunk, and tail, and BBox centroid displacement. We evaluated the method using the Sweetfish dataset from Fish Tracking Challenge 2024 under conditions where multiple pseudo-detection losses were applied to detection sequences derived from manual annotation. The results showed that the proposed method outperformed OC-SORT in all conditions, achieving an ID consistency of 1.0 and zero ID switches (IDSW = 0), particularly in frames where detections were available. This demonstrates that the proposed method can effectively suppress ID switches inherent to online tracking within a range of practically acceptable latency.

January 22 (Thursday), 13:00-14:30

B1 Meeting Room 4

OS21 AROB: System Sensing and Its Applications I

Organizer: Kent Nagumo (Aoyama Gakuin Universiy, Japan)
Hironobu Sato (Kanto Gakuin University, Japan)

OS21-1 Improved DFA for Fluctuation Analysis of Short Time Series Data

Ryouma Saiki[1] and Masafumi Uchida[2]
([1]The University of Electro-Comunications, Graduate School of Informatics and Engineering, Japan)
([2]The University of Electro-Comunications, Japan)

Data acquired from time variation of events are referred to as "time series". Time series data exhibit a unique “fluctuation”. The nature of fluctuation changes depending on the event. Thus, analysis of fluctuation is useful for examining differences in events. In previous research, the fluctuation analysis was applied to the data which were obtained from pick and place action. By examining the fluctuation, the study attempted to assess proficiency in that action. The study used detrended fluctuation analysis (DFA). However, DFA is originally suitable for long-length data and the data used in previous research were too short. Consequently, the issue of decreased research precision is significant. Because of this, this study investigated a modified DFA suitable for aptitude for analyzing short-length data. To evaluate this, the method’s effect on maintaining DFA precision when analyzing short-length data were assessed by comparing different-length pattern’s results derived from the same original dataset. The results showed differences across conditions, making it difficult to clearly validate the new DFA. However, the new DFA has notable effects of the suppression of standard deviation.

OS21-2 Effects of Vibrotactile Apparent Movement of the Neck on Postural Sway

Keisuke Sakuma and Masafumi Uchida
(The University of Electro-Comunications, Graduate School of Informatics and Engineering, Japan)

In this study, vibrotactile stimuli utilizing apparent movement were applied to the human neck, and the center of pressure (COP) during standing posture was measured and analyzed. In previous studies, experimenters selected the parameters of apparent movement and presented them to participants. However, apparent movement is a type of illusion that depends on individual subjective perception. Therefore, it is desirable that the parameters be finely adjustable and freely selectable by the participants themselves. To achieve this, a new device was developed that allows participants to select the parameters of apparent movement on their own. As a result, it was observed that vibrotactile stimulation using forward and backward apparent movement selected by the participants tended to induce anterior–posterior postural sway (COP). This finding suggests the possibility of presenting more intuitive directional sensations independent of conscious intention. Such technology is expected to facilitate applications such

as hazard warnings, sports instruction, and the transmission of information—such as slope or temperature—that cannot be perceived through vision or hearing.

OS21-3 Effects of Vibrotactile Stimulation to the Hand on Tremor

Koki Okamoto and Masafumi Uchida
(The University of Electro-Communications, Graduate School of Informatics and Engineering, Japan)

Physiological tremor is an involuntary movement commonly observed in healthy individuals, often becoming noticeable during states of tension or when performing delicate tasks. To enable more accurate actions in daily life, there is a need for an easily wearable tremor suppression device. This study evaluated the effects of presenting various vibrotactile stimulation to the hand on physiological tremor. It was hypothesized that vibrotactile stimulation introduces disturbances to the peripheral sensorimotor system, thereby influencing the tremor generation mechanism. The vibration parameters investigated were the input voltage and the vibration duration . A higher input voltage corresponded to a stronger vibration intensity. The successful identification of vibration parameters that suppress physiological tremor will directly contribute to the development of compact and easily wearable tremor suppression devices. As a result, it was found that vibrotactile stimulation influenced physiological tremor, although the tendency for amplification or suppression was mixed in the time domain following the cessation of the stimulus and varied among subjects.

OS21-4 Effects of Air Environment on Students' Behavior and Psychophysiological States

Keiji Osumi[1], Koya Hiraishi[1], Shu Kobayashi[2], Nagul Cooharojananone[3], Hirokazu Doi[1] and Kosuke Oiwa[1]
([1]Nagaoka University of Technology, Japan)
([2]Takasago Thermal Engineering Co.,Ltd, Thailand)
([3]Chulalongkorn University, Thailand)

It is widely recognized that the air environment significantly affects human psychological and physiological states. However, most previous studies on the air environment and learners' psychological states have relied on assessment, and research quantitatively examining the relationship between the air environment and psychophysiological responses remains limited. The purpose of this study is to clarify the effects of temperature and humidity conditions in a study room on students' behavior and psychophysiological states. Eleven university students performed self-paced study tasks for one hour under different environmental conditions, during which heart rate variability (HRV) and subjective comfort were measured. The results revealed a discrepancy between physiological and psychological responses. Specifically, despite the suppression of parasympathetic nervous activity associated with increased temperature, participants reported improved subjective comfort. Furthermore, a non-linear relationship between temperature and pulse wave intervals was observed in approximately half of the participants, highlighting the existence of significant individual differences that cannot be captured by simple linear models. These findings suggest the necessity of considering individual differences that escape linear modeling when optimizing learning environments.

OS21-5 Trial-by-Trial Variation in State Empathy for Video Stimuli: Scale Development and Validation of Variation Patterns

Miho Iijima and Shogo Matsuno
(The University of Electro-Communications, Japan)

While empathy research has traditionally focused on trait-level individual differences, practical applications in medical education, bullying prevention, and media production require understanding what stimuli elicit empathy. This study developed an 8-item state empathy scale measuring cognitive empathy (CE) and affective empathy (AE) for video stimuli. Three participants viewed 24 emotional videos (72 trials total). Results showed that AE demonstrated good reliability (α = .831) and strong convergent validity with self-other overlap (IOS: r = .845) and emotional congruence (EC: r = .240). CE showed acceptable reliability (α = .623) and predicted negative correlation with response time (r = -.257). Critically, variance decomposition revealed that 84% of empathy variation was stimulus-driven, with only 16% attributable to individual differences. Emotion category analysis showed sadness videos elicited the strongest empathy (AE: M = 5.56), with large effect size compared to joy videos (d = 0.98). These findings support trial-level research approaches and provide foundational knowledge for designing empathy-eliciting educational materials and media content.

OS21-6 The Attempt to Reproduce an Image Impression Survey using the SD Method with CLIP

Atsushi Shibata
(Department of Informatics, the University of Electro-Communications, Japan)

This study investigates the extent to which large-scale Vision-Language Models (VLMs) have acquired a human-like "Kansei semantic space." While VLMs like CLIP and SigLIP share a common embedding space for images and text, it remains unclear whether they capture abstract impressions (e.g., "softness" or "warmth") derived from visual features, akin to human cross-modal perception. To verify this, we conducted a comparative analysis between text-only models (Word2Vec, BERT) and VLMs using the Semantic Differential (SD) method. We utilized a dataset of controlled synthetic images (simple shapes and colors) and antonym pairs representing abstract impressions. By applying Z-score normalization to mitigate embedding anisotropy, we quantified the association between visual primitives and impression words. Our results demonstrate that VLMs structure concepts based on visual co-occurrence, distinct from the text-based statistics of Word2Vec and BERT. Notably, VLMs successfully inferred abstract impressions—such as associating spheres with "Soft" and warm colors with "Active"—without explicit textual descriptions. These findings align with psychological insights into human perception, suggesting that VLMs possess a latent capacity for affective judgment and can function as agents capable of understanding subjective impressions.

January 22 (Thursday), 14:45-16:00

3F Meeting Room 32

OS11 AROB: Control of Surface and Underwater Robots

Organizers: Shinichi Sagara (Kyushu Institute of Technology, Japan)
Yuta Hanazawa (Kyushu Institute of Technology, Japan)

OS11-1 Experiments on a 3-link dual-arm underwater robot using a resolved acceleration control for UVMS - Application of fictitious reference iterative tuning method -

Naoya Kakihara[1], Yuta Hanazawa[1], Shinichi Sagara[1] and Radzi Ambar[2]
([1]Kyushu Institute of Technology, Japan)
([2]Universiti Tun Hussein Onn Malaysia, Malaysia)

In recent years, research and development have been devoted to controlling underwater robots equipped with manipulators (known as Underwater Vehicle-Manipulator Systems, or UVMS) that can perform underwater tasks in place of humans. We have proposed a Resolved Acceleration Control (RAC) method, a position control approach for UVMS, and have validated its usefulness through experiments using a free-floating, 3-link dual-arm underwater robot. To maximize the performance of the RAC method, control parameters including feedback gains must be properly tuned. However, this tuning has traditionally been done through manual trial and error, relying on the operator's experience, which makes the adjustment process time-consuming and costly. To overcome this problem, we apply the Fictitious Reference Iterative Tuning (FRIT) method to the RAC system. FRIT allows for offline parameter optimization using only a single set of closed-loop experimental data without relying on an accurate mathematical model of the system. In this paper, the effectiveness of the tuning method for the underwater robot affected by nonlinear hydrodynamic forces is demonstrated through experiments comparing the control performance before and after optimization.

OS11-2 Force control experiment of a 3-link dual-arm underwater robot using a force sensor unit with waterproof foil strain gauges

Yuki Moritaka[1], Kensho Osugi[1], Yuta Hanazawa[1], Shinichi Sagara[1] and Radzi Ambar[2]
([1]Kyushu Institute of Technology, Japan)
([2]Universiti Tun Hussein Onn Malaysia, Malaysia)

The underwater environment is an extreme condition for humans, making direct underwater operations hazardous. Consequently, the development of underwater robots equipped with manipulators (UVMS: Underwater Vehicle-Manipulator System) to perform tasks in place of humans has been pursued. We have developed position control methods for UVMS and have verified their effectiveness through simulations and experiments. Force control is also necessary for UVMS, and we have conducted force control experiments using a force sensor with a waterproof mechanism. To conduct various force control experiments of UVMS in the future, it is considered essential to miniaturize the underwater force sensor unit. In this paper, we aim to develop a small, waterproof force sensor unit using strain gauges and verify its usefulness through basic experiments. First, a preliminary ground experiment was conducted under a constant load to understand basic output characteristics. Next, we placed the prototype units at both the base of the manipulator hand and the fixed underwater object (wall), and conducted a pushing experiment using a constant force to verify the sensor's usefulness.

OS11-3 Continuous experiments from fully floating position control to semi-fixed force control of a 3-Link dual-arm underwater robot using RAC method for UVMS

Kensho Osugi[1], Yuta Hanazawa[1], Shinichi Sagara[1] and Radzi Ambar[2]
([1]Kyushu Institute of Technology, Japan)
([2]Universiti Tun Hussein Onn Malaysia, Malaysia)

In recent years, various ocean development activities such as offshore oil and natural gas extraction, renewable energy development, seabed mineral resource exploitation, and marine spatial utilization have been carried out. Advances in marine science and technology, especially in underwater robotic systems, are essential to promoting these activities. These systems, which include advanced ocean mobility platforms such as the Autonomous Underwater Vehicles (AUVs) and the Remotely Operated Vehicles (ROVs), collectively represent a critical and indispensable foundational technology for the future of ocean science and engineering. Furthermore, among the previously mentioned AUVs and ROVs, research and development are also being conducted on Underwater Vehicle-Manipulator Systems (UVMSs), designed for tasks such as manipulating underwater objects. We have proposed Resolved Acceleration Control (RAC) methods for UVMS, as a position control. We also have proposed a force control method, which combines the RAC methods with position-based impedance control. The effectiveness of these methods has been verified through simulations and experiments using a three-link dual-arm underwater robot. However, to perform force control on an object, the UVMS has to first grasp it. In this paper, we demonstrate the feasibility of these tasks through continuous experiments applying the proposed method to a three-link dual-arm underwater robot.

OS11-4 Design of a stable impedance controller for a manipulator attached to a surface vehicle

Yuichiro Taira and Toyoyuki Honjo
(National Defense Academy, Japan)

This paper deals with the design of a motion and force controller based on the impedance control strategy for a manipulator (a robotic arm) that is attached to a surface vehicle (e.g., a boat). The authors have devised an impedance control scheme in view of the dynamics of the vehicle whose position and orientation fluctuate due to environmental disturbances such as water waves. However, the stability of the control system is not guaranteed, and furthermore the mechanical impedance is directly affected by the disturbances. In this paper, we develop an impedance controller with a disturbance estimator. Its features are (1) to theoretically ensure the stability of the overall control system, (2) to attenuate the influence of the disturbances on the mechanical impedance by including the estimator in the impedance controller, and (3) to enable the improvement of the estimated performance by increasing the value of a design parameter concerning the disturbance estimator. Moreover, the theoretical results were supported by those of numerical simulations for a two-link manipulator mounted on a surface vehicle, which can move only in a horizontal plane.

OS11-5 Design of Small and Lightweight Excavation ROV for archaeological survey

Takahiro Yamamura, Tatsumu Matsumae and Norimitsu Sakagami
(Ryukoku University, Japan)

In recent years, underwater robotic vehicles have increasingly replaced human divers in archaeological surveys. Autonomous Underwater Vehicles (AUVs) are useful tools for large-scale visual survey applications. However, Remotely Operated Vehicles (ROVs) are deployed for delicate intervention tasks such as excavation. According to UNESCO, the maximal preservation of underwater sites and artifacts is required to protect them from destruction. Therefore, ROVs remain indispensable tools for excavation because of their real-time intervention capabilities. Large ROVs with excavation tools are typically used to handle archaeological objects during archaeological intervention. Consequently, their operation requires large vessels, which increases the operational costs. Nevertheless, to date, there have been few reports of excavations conducted using small and lightweight ROVs other than traditional excavation tools. In this study, we aimed to develop a small and lightweight ROV equipped with an underwater excavation device to enable low-cost and easily deployable archaeological excavation.

January 22 (Thursday), 14:45-16:30

3F Meeting Room 33

OS14 AROB: Intuitive Human-System Interaction II

Organizer: Masao Yokota (Fukuoka Institute of Technology, Japan)

OS14-1 Experiments of Multiple IoT Sensors for Early Heavy Rainfall Warning System with Geographical and Micrometeorological Factors

Noriki Uchida[1], Tomoyuki Ishida[1] and Yoshitaka Shibata[2]
([1]Fukuoka Institute of Technology, Japan)
([2]Iwate Prefectural University, Japan)

In recent years, sudden heavy rainfall has caused severe flooding, property damage, and loss of life worldwide. Traditional forecasting systems, such as AMeDAS and radar-based methods, operate at kilometer-scale resolution, making it difficult to predict highly localized weather phenomena. Therefore, this study proposes the Early Warning System for Heavy Rainfall, which integrates micrometeorological forecasts from cloud-based services with detailed geographical data collected by numerous IoT sensors. In the proposed system, atmospheric changes are detected using real-time micrometeorological prediction services via Web APIs, while numerous IoT sensors monitor local conditions such as 3 axis accelerometer, soil moisture, flooding, temperature, and structural shifts within small areas. To improve prediction accuracy, the system proposes the Enhanced MQTT method to increase the IoT network connections and the Kalman Filter-based anomaly detection model that combines time-series atmospheric and geographical data, continuously adjusting estimates and calculating abnormality scores using Gaussian distribution models. This paper especially introduces the implementation of IoT devices to sense the flow of water or ground, and the ongoing field experiments were reported for the future works.

OS14-2 Case study on the attempting 3D data creation from commercially available point cloud data

Kodai Tsushima
(EBARA Corporation, Japan)

Advanced simulation techniques based on three-dimensional spatial data have become increasingly important in construction informatics. Point cloud data provide high-fidelity representations of as-built environments; however, their extremely large data volume results in substantial computational overhead, making real-time collision-aware simulation difficult on standard PC platforms. This study presents a data reduction and simulation workflow that enables immersive VR-based simulation by converting point cloud data into a lightweight 3D model. Point cloud data captured using a Matterport Pro3 camera were first converted into an IFC-format model and then transformed into STEP format with texture information for compatibility with a VR simulation platform. This pipeline reduced the data size from 7.5 GB to 79.6 MB while preserving geometric features necessary for simulation. The accuracy of the generated 3D model was evaluated by comparing distance measurements with those obtained from point-cloud-based simulations. The maximum observed deviation was approximately 62 mm, which is acceptable relative to the operational tolerance of crane-based equipment handling. The results indicate that the proposed approach significantly reduces computational requirements while maintaining sufficient accuracy for collision detection and transport feasibility analysis in immersive simulation environments.

OS14-3 Pipe Point Cloud Recognition Using BIM Data from Fukushima Daiichi Nuclear Power Plant and Its Applications

Akio Doi[1], Toru Kato[2], Hiroki Takahashi[2], Meguru Yamashita[2] and Gao Zhi Yi[3]
([1]Iwate Prefectural University, Japan)
([2]Freelancer, Japan)
([3]Shanxi Vocational University of Engineering Science and Technology, China)

We are developing advanced 3D digital modeling and simulation technologies aimed at reducing radiation dose rates at the Fukushima Daiichi Nuclear Power Plant Unit 3 (1F3). To achieve this, we have established a methodology for estimating dose rates by acquiring high-density point cloud data with a 3D laser scanner, followed by structural recognition and modeling of key components. However, due to the physical constraints of the plant environment, the 3D laser scanner cannot capture sufficient point cloud data for narrow, elevated piping—commonly referred to as “macaroni piping”—which limits the accuracy of detailed modeling. To address this limitation, the present study utilizes Building Information Modeling (BIM) data of the piping system derived from 1F3 design drawings. Using these BIM data, we constructed a learning model based on the PointNeXt architecture to classify piping components into three categories: pipe bodies, fittings, and valves.

OS14-4 Development of a Vision-Based Intelligent Recognition System for Patrol Inspection Robots

Gao Zhi Yi[1], Akio Doi[2], Toru Kato[3], Hiroki Takahashi[3] and Meguru Yamashita[3]
([1]Shanxi University of Engineering Science and Technology, China)
([2]Iwate Prefectural University, Japan)
([3]Freelancer, Japan)

This work presents a vision-based intelligent recognition system developed as part of ongoing research on industrial visual inspection and automated monitoring. The system targets unmanned inspection rooms in legacy industrial facilities, where equipment typically lacks modern data transmission capabilities. Using a YOLO-based detection framework, the system classifies devices into digital displays, indicator lights, and pointer-type meters, followed by secondary processing to extract numerical values or status information from camera images. This approach provides a practical alternative to sensor-based data collection and supports automated reading and anomaly detection under varying environmental conditions. The study also contributes to broader efforts in intelligent inspection, meter-reading automation, and the integration of computer vision algorithms into mobile patrol inspection robots. The proposed method demonstrates high accuracy and robustness, forming a foundation for future deployment in real-world industrial monitoring applications.

OS14-5 Collaborative Mixed Reality System for Traditional Craft Digital Archives

Shion Kimura and Tomoyuki Ishida
(Fukuoka Institute of Technology, Japan)

Japan’s traditional crafts, which are supported by the advanced skills of its artisans, have a long history. To pass these crafts on to future generations, it is important to show their appeal to young people both in Japan and abroad and generate new interest. In our previous study, we developed virtual /augmented reality systems for experiencing Japanese culture. However, these systems could not be shared by multiple users simultaneously or used remotely. To overcome the experiential isolation and geographical constraints, we develop a system that allows multiple users to collaboratively view craft digital archives in the same mixed reality space. This system will enable people from around the world to participate, promoting international cultural exchange via real-time discussions while viewing traditional Japanese crafts. To evaluate the proposed system, we conducted an evaluation experiment involving six university students. All participants reported that the system increased their interest in traditional crafts; many of them stated that it improved their understanding. However, some pointed out a lack of historical and cultural information. The interaction of the system with remote users and gesture functionality was generally well-received, demonstrating its effectiveness in promoting cultural understanding and remote communication.

OS14-6 Proposal of a Virtual and Real Space Integrated Tourism Support System

Kazuki Beppu and Tomoyuki Ishida
(Fukuoka Institute of Technology, Japan)

In recent years, the use of digital technology in tourism has increased, and virtual reality (VR) systems, which allow users to explore tourist destinations in a simulated environment, have attracted attention. However, traditional VR tourism solutions lack integration with actual tourist behavior and mechanisms for applying acquired information to travel plans and on-site activities. Similarly, although research on location-based navigation and augmented reality guides has progressed, most studies focus on on-site use, with limited efforts to establish bidirectional links between virtual and real spaces. Therefore, this study developed a metaverse-based pretourism experience system that aligns location information with VR space and seamlessly connects the two. Users can explore the topography and scenery of tourist destinations in VR space and assign check tags to each spot based on real-world locations, allowing them to incorporate their virtual experiences into on-site travel plans. Ultimately, this system aims to make pretourism experiences directly contribute to tourism decision-making, resulting in enhanced planning efficiency and improved tourism experience value.

OS14-7 Multi-Information Sharing Environment using Mixed Reality Technology

Yuri Tanaka and Tomoyuki Ishida
(Fukuoka Institute of Technology, Japan)

Mixed reality (MR) is currently applied in various fields, including medicine and education. MR technology enables the superimposition of digital content onto real space, allowing users to experience diverse content in a three-dimensional, interactive manner. However, many MR systems use complex operational methods, resulting in a high learning cost. Therefore, this study used MR technology to develop a multi-information sharing environment for multiple users in the same space, featuring a simple operating system that beginners can use intuitively. In designing and implementing this multiuser MR system, we created an interface based on simple button operations. The system is user-centered and comprises a multi-information sharing system, a material presentation module, an artificial intelligence assistant module, cloud storage, an audio recognition engine, a text summarization engine, and a Photon server. The proposed system reduces the user fs operational burden and supports smooth face-to-face communication through spatial and visual information sharing. This enables a natural and collaborative information sharing experience in the MR environment.

January 22 (Thursday), 14:45-16:00

B1 Meeting Room 1

OS16 AROB: Robotics with Intelligence and/or Informatics II

Organizers: Tetsuya Kinugasa (Kindai University, Japan)
Masatoshi Hatano (Nihon University, Japan)

OS16-1 Implementation of the Diadema setosum Eradication System for a Small Underwater Robot

Toma Mori, Yamato Kayahara, Bingyao Lu, Zhaohui Tan, Masanori Sato and Kazuhiko Matsuoka
(Nagasaki Institute of Applied Science, Japan)

This paper proposes an autonomous eradication system for Diadema setosum using a small underwater robot to improve the "Isoyake" (seafloor desertification) phenomenon. The system integrates real-time object recognition with an autonomous approach control algorithm for crusher-equipped manipulators. For target detection, we developed a model using YOLO11 trained on datasets from actual sea areas, achieving a mean Average Precision (mAP) of 0.848. To ensure edge-side performance, the model was optimized for a Raspberry Pi 5, attaining a processing speed of approximately 10 FPS at a resolution of 320x240 pixels, which provides sufficient real-time capability for underwater operations. Furthermore, an autonomous approach algorithm based on visual servoing was implemented, utilizing the center coordinates and bounding box size of the detected target to guide the robot into the manipulator's workspace. Experimental results demonstrate the feasibility of the proposed recognition and control framework, laying the foundation for fully autonomous sea urchin eradication in marine environments.

OS16-2 Reconstructing Dinosaur Hindlimb Locomotor Motion Using a Bioinspired Musculoskeletal Robot Based on Crocodilian Muscle–Tendon Structure

Evan Kojima[1], Kazuki Ito[1], Ryosuke Nawama[1], Tsukasa Okoshi[2], Kentaro Chiba[2], Damdinsuren Idersaikhan[4], Yasuhiro Sugimoto[1] and Tetsuya Kinugasa[3]
([1]Department of Mechanical Engineering, The University of Osaka, Japan)
([2]Department of Biosphere-Geosphere Science, Okayama University of Science, Japan)
([3]Department of Mechanical Engineering, Kindai University, Japan)
([4]Department of Institute of Paleontology, Mongolian Academy of Sciences, Mongolia)

This study investigates the hindlimb locomotor mechanism of the non-avian dinosaur Protoceratops andrewsi using a bioinspired musculoskeletal robot. The hindlimb musculature was inferred from anatomical observations of the extant crocodilian Crocodylus porosus and implemented on a physical model reconstructed from CT and photogrammetry data. Active muscle functions were reproduced using pneumatic actuators, and muscle activity sequences were designed based on crocodilian electromyography patterns. Robotic experiments demonstrated that the reconstructed hindlimb reproduced a continuous stance-to-swing transition, indicating that the inferred musculature and activation patterns are sufficient to generate walking motion in a physical environment. These findings highlight the value of musculoskeletal robots for experimentally evaluating locomotor hypotheses in extinct animals and contribute to understanding functional principles underlying dinosaur locomotion.

OS16-3 Basic Study on Tomato Harvesting with Dual-Arm Cooperation

Tatsuya Horikawa[1], Toma Mori[1], Huizhen Li[1], Zhuoyu Zeng[1], Zhaohui Tan[1], Masanori Sato[1] and Satoshi Makita[2]
([1]Department of Electronics and information Technology,
Graduate School of Nagasaki Institute of Applied Science, Japan)
([2]Department of Intelligent Mechanical Engineering,
Fukuoka Institute of Technology, Japan)

In this study, we implement Dual-arm tomato harvesting robot to address the limitations of conventional methods, which rely on color-based detection and Single-arm harvesting. Regarding the harvesting method, conventional Single-arm lateral harvesting often results in collisions with fruits, making grasping difficult. Approaching from below can mitigate this issue, but fully ripe fruits are typically located at the top, complicating their harvest. To solve this problem, we implement Dual-arm harvesting: one manipulator approaches from below to secure the fruit from behind, while the other manipulator harvests from the front side, ensuring reliable collection of fully ripe tomatoes. By integrating Dual-arm harvesting, the proposed system enhances both the accuracy and reliability of tomato harvesting. The effectiveness of the proposed system is evaluated in a simulated environment, and validation in an actual cultivation environment is planned for future work. Additionally, we plan to implement AI-based fruit detection to improve the accuracy and reliability of tomato harvesting in the future.

OS16-4 A Deep Learning-Based Framework for Bed-Leaving Detection System Using YOLO and Siamese Network

Zhaohui Tan[1], Jiaqi Xu[1], Masanori Sato[1], Manabu Yamaji[1] and Takeshi Ikeda[2]
([1]Nagasaki Institute of Applied Science, Japan)
([2]Sanyo-Onoda City University, Japan)

This study proposes a non-contact bed-leaving detection system using depth images for nighttime monitoring in nursing care facilities. In our previous work, YOLO:You Only Look One was used as an object detection method, where the human body was divided into multiple posture-related classes, allowing multiple posture detections within a single frame. In this paper, YOLO is simplified to detect only the human body region, and a Siamese network which is one of the image classfication method evaluates similarity between the detected target human and a reference posture set to identify bed-leaving behavior. Experimental results demonstrate that the proposed framework achieves high accuracy and robustness in human detection and posture classification.

OS16-5 Improving Hand Detection Accuracy Using Machine Learning in Assistive Devices for the Visually Impaired

Takeshi Ikeda[1], Masahiro Shiotani[1], Yoshiki Tanaka[2], Fusaomi Nagata[1], Masanori Sato[3] and Seiji Furuno[2]
([1]Sanyo-Onoda City Unviersity, Japan)
([2]National Institute of Technology, Kitakyushu College, Japan)
([3]Nagasaki Institute of Applied Science, Japan)

There are approximately 310,000 visually impaired people in Japan. Research and developments are being conducted to provide the necessary support for visually impaired people in their daily lives. Our laboratory has been researching an interactive guidance system to support visually impaired people, focusing on assisting them in using tools. This system focuses on the movements of visually impaired people as they search for and grasp the tools they want to use in their daily lives. It uses image processing to detect the tool using a camera and guides the user's hand to the tool. However, if the orientation of the hand changes during guidance, the hand cannot be detected, and situations where guidance is not possible often occur. In this study, we aim to address this issue by utilizing machine learning to detect the hand that is the target of guidance, creating a detection model that adapts to changes in hand posture during guidance, thereby reducing situations where the hand cannot be detected and guidance instructions cannot be issued.

January 22 (Thursday), 14:45-15:30

B1 Meeting Room 2

OS18 AROB: Swarm and Modular Robotics II

Organizer: Ryusuke Fujisawa (The University of Kitakyushu, Japan)

OS18-1 Experimental Verification of a Health Monitoring Scheme for the Connection Sequence of Modular Robots

Yuki Takagi[1], Kazuki Shibata[2], Hiroshi Oku[1], Yuki Tanigaki[1], Guang Yang[1], Xixun Wang[1], Ryohei Michikawa[2] and Fumitoshi Matsuno[1]
([1]Department of Electronics and Information System Engineering, Osaka Institute of Technology, Japan)
([2]Department of Mechanical Engineering and Science, Kyoto University, Japan)

Our goal is to develop a health monitoring framework for the connection sequence between modules in an autonomous modular robot designed for lunar exploration and base construction. The robot system under development consists of three types of modules: limb modules, body modules, and end-effector modules, all equipped with genderless connection mechanisms. In our scenario, each module is launched to the lunar surface fixed to a pallet and assembled into the desired configuration on-site. Our scenario is that, after landing on the lunar surface, a limb module placed on a shipping container pallet begins to connect to an end-effector module on the pallet, forming a robotic system. Our health monitoring system monitors the module connection sequence in four steps. Especially in this paper, we will focus on the step where the tip of the limb is aligned with the connector part of the end effector, by moving limb’s joints. Our health monitoring system automatically observes whether a camera attached at the tip of the limb successfully detects the AR-marker at the center of the connector part of the end effector or not. We implement the health monitoring scheme into our real modular robot in order to demonstrate usefulness of our framework.

OS18-2 Analyzing Robustness and Performance Trade-offs in Modular Robot Task Allocation Using a Two-Stage Bottom-Up Optimization Framework

Yuki Tanigaki[1], Mikito Yano[1], Tomohiro Hayakawa[2], Ryusuke Fujisawa[3] and Fumitoshi Matsuno[1]
([1]Department of Electronics and Information Systems Engineering,
Osaka Institute of Technology, Japan)
([2]Department of Mechanical Engineering, Faculty of Engineering, Shizuoka University, Japan)
([3]Faculty of Environmental Engineering, Kitakyushu University, Japan)

The performance of a modular robot varies significantly depending on its configuration, making task allocation a complex problem. To address the interdependence between robot configuration and task allocation, this study focuses on a two-stage bottom-up optimization framework. In the first stage, Pareto-optimal robot configurations are explored through multi-objective optimization. In the second stage, several representative configurations are selected from the obtained Pareto front, and task scheduling optimization is performed for each selected configuration. The framework is applied to a lunar base construction scenario, where task allocation, module distribution, robot performance, and failure probabilities are systematically varied. Through this analysis, we investigate how the balance between robot performance and required numbers of modules affects task execution efficiency under failure-prone conditions.

OS18-3 Automatic Self-Assembly of Modular Robots: Experimental Demonstration of a Two-Wheeled Rover Formed from Limb, Wheel, and Tail Modules

Guang Yang[1], Ryohei Michikawa[2], Yuto Fukao[2], Takashi Takuma[1], Masahiro Ikeda[3], Xixun Wang[1],
Yuya Shimizu[4], Hidefumi Imamura[1], Yuta Shimizu[1], Yoshiki Higashino[1], Haruho Mitsunaga[1],
Hazuki Nunome[1], Yosuke Izumino[1], Ryosuke Fujisawa[5], Tetsushi Kamegawa[4] and Fumitoshi Matsuno[1]
([1]Osaka Institute of Technology, Japan)
([2]Kyoto University, Japan)
([3]Kindai University, Japan)
([4]Okayama University, Japan)
([5]The University of Kitakyushu, Japan)

Modular robots for lunar base construction and planetary exploration offer high portability and reconfigurability; however, they must also possess the capability of autonomous self-assembly to form functional configurations without human intervention. Such a capability is essential not only for improving transportation efficiency but also for enabling adaptive morphological reconfiguration according to environmental conditions and task requirements. This paper presents an experimental demonstration of automatic self-assembly using real modular robot hardware. A two-wheeled rover is autonomously assembled from one limb module, two wheel modules, and one tail module through coordinated motions of the limb module. The assembly process includes mechanical docking, initiation of wheel-driven locomotion, and descent from a pallet to the ground. Locomotion experiments conducted on sandy terrain confirm that the assembled rover achieves sufficient traction and posture stability, demonstrating a smooth transition from self-assembly to mobile operation. The results provide fundamental insights into practical self-assembly of modular robots and serve as a basis for future extensions to more complex module configurations and lunar-like environments.

January 22 (Thursday), 14:45-16:00

B1 Meeting Room 3

OS20 AROB: Swarm-Machine Interaction: Models and Foundations for Collective Behavior II

Organizers: Hiroaki Kawashima (University of Hyogo, Japan)
Takayuki Niizato (University of Tsukuba, Japan)
Hitoshi Habe (Kindai University, Japan)

OS20-1 Data-Driven Control of a Magnetically Actuated Fish-Like Robot

Akiyuki Koyama and Hiroaki Kawashima
(Graduated School of Information Science, University of Hyogo, Japan)

Magnetically actuated fish-like robots offer promising solutions for underwater exploration due to their miniaturization and agility; however, precise control remains a significant challenge because of nonlinear fluid dynamics, flexible fin hysteresis, and the variable-duration control steps inherent to the actuation mechanism. This paper proposes a comprehensive data-driven control framework to address these complexities without relying on analytical modeling. Our methodology comprises three core components: 1) developing a forward dynamics model (FDM) using a neural network trained on real-world experimental data to capture state transitions under varying time steps; 2) integrating this FDM into a gradient-based model predictive Control (G-MPC) architecture to optimize control inputs for path following; and 3) applying imitation learning to approximate the G-MPC policy, thereby reducing the computational cost for real-time implementation. We validate the approach through simulations utilizing the identified dynamics model. The results demonstrate that the G-MPC framework achieves accurate path convergence with minimal root mean square error (RMSE), and the imitation learning controller (ILC) effectively replicates this performance. This study highlights the potential of data-driven control strategies for the precise navigation of miniature, fish-like soft robots.

OS20-2 Modeling and Analysis of Fish Interaction Networks under Projected Visual Stimuli

Hiroaki Kawashima, Raj Rajeshwar Malinda and Saeko Takizawa
(University of Hyogo, Japan)

This paper addresses the estimation of a dynamic interaction network, a network of influence among individuals, under projected visual stimuli to quantify the influences of inter-individual interactions and external stimuli on collective behavior. Building upon our previously proposed network estimation model, which assumes a Boids-type model and employs a sparse regression framework to infer inter-individual influence networks from trajectory data, we extend the formulation by introducing a stimulus term. This enables the model to capture how individuals react to and propagate externally projected visual stimuli within the group. The resulting framework allows simultaneous estimation of inter-individual and stimulus-related interaction strengths. We also introduces entropy-based indices to capture the possible biases of individuals' influence. Our experiments with fish schools under projector-based visual stimuli demonstrate the effectiveness of the proposed indices in quantifying schooling behavior and identifying influential individuals within the group, serving as the basis for real-time, interpretable metrics of collective dynamics.

OS20-3 Tail-Beat Synchronization of Plecoglossus altivelis under Different Light Wavelengths

Kohei Ohashi[1], Shoma Kamata[1], Funo Suzuki[1], Taiki Nishimura[1], Hikaru Yamano[1], Hiroaki Kawashima[2], Hitoshi Habe[3] and Takayuki Niizato[1]
([1]Institute of Systems and Information Engineering, University of Tsukuba, Japan)
([2]Graduate School of Information Science, University of Hyogo, Japan)
([3]Faculty of Informatics, Kindai University, Japan)

Schooling may be guided by non-invasive visual stimuli, yet mechanisms linking light environments to collective structure remain unclear. We tested whether light wavelength reorganizes pair behavior in juvenile Plecoglossus altivelis through propulsion-phase coordination. Two-fish pairs (nine datasets per wavelength) swam freely in a 2 m × 2 m arena under blue, red, and white light in randomized order, recorded with infrared imaging. A custom YOLOv11 pose model (six keypoints) enabled time-resolved tail-beat extraction even during partial overlap/occlusion. Tail-tip lateral displacement relative to the instantaneous body axis was filtered, and synchrony was quantified by windowed Pearson correlation, classifying windows as in-phase, anti-phase, or non-correlated. Blue light yielded more cohesive and polarized motion than red, whereas state proportions changed little across wavelengths. Conditioning on synchronized windows, anti-phase occurred more often than in-phase. This anti-phase dominance under free swimming supports a multi-layer view of schooling in which low-level propulsion-phase interactions, potentially hydrodynamic and/or flow-sensing, coexist with higher-level alignment dynamics.

OS20-4 Self-Image Projection Effect on the Collective Behavior of Plecoglossus altivelis

Funo Suzuki[1], Kohei Ohashi[1], Shoma Kamata[1], Taiki Nishimura[1], Hikaru Yamano[1], Hiroaki Kawashima[2], Hitoshi Habe[3] and Takayuki Niizato[1]
([1]Institute of Systems and Information Engineering, University of Tsukuba, Japan)
([2]Graduate School of Information Science, University of Hyogo, Japan)
([3]Faculty of Informatics, Kindai University, Japan)

Visual information plays an extremely important role in the collective behavior of fish, and biologically plausible motion in particular serves as a crucial cue for conspecific recognition. However, conventional light-induced guidance has largely relied on the one-sided presentation of pre-programmed movements, making flexible, situation-specific intervention difficult. In this study, we developed a method for Plecoglossus altivelis(P. altivelis) that projects manipulated self-shadows, namely silhouettes derived from the fish's own behavioral history with either a time delay or a flipped, and examined whether such visual stimulation can modulate group behavior. For analysis, based on the inter-individual distance between two fish and their distance to the tank walls, we classified the pair state into three categories: collective wall-following (anxiety-like), collective exploration (sociability-like), and separation. We found that interaction with the manipulated self-shadow significantly reduced thigmotaxis while maintaining cohesion and promoted exploratory swimming in the central region of the tank. These results suggest that a fish's "past self-shadow" was accepted as a social partner for interaction, demonstrating that motion-based visual information alone can guide and activate fish groups without disrupting group cohesion.

OS20-5 Speech Structure Through the Lens of Acoustic Feature Interactions: A Comparative Analysis of Natural and Synthetic Speech with Integrated Information Theory

Daichi Tasaki and Takayuki Niizato
(Institute of Systems and Information Engineering, University of Tsukuba, Japan)

The rapid integration of synthetic speech into daily life makes understanding its cognitive impact urgent. Traditional subjective evaluations regarding the "Uncanny Valley" often yield conflicting findings. To address this issue, we introduce an objective framework using Integrated Information Theory (IIT) to quantify the information structure of acoustic features. We analyzed formant frequencies and intensity in natural speech from 11 participants and 10 synthetic voices. We calculated "integrated information" (Φ) and the "Main Complex" to quantify structural interdependence. Additionally, we used Earth Mover's Distance (EMD) to compare structural similarities. Results indicated that synthetic speech exhibits wider structural variability than natural speech. Notably, some synthetic voices structurally close to the natural average did not align with high subjective naturalness. This mismatch suggests listeners might still perceive such voices as "mechanical," pointing toward the "Uncanny Valley." This study establishes IIT as a powerful tool for speech analysis. This framework paves the way for future studies correlating objective metrics with subjective evaluations.

January 22 (Thursday), 14:45-16:00

B1 Meeting Room 4

OS22 AROB: System Sensing and Its Applications II

Organizers: Atsushi Shibata (The University of Electro-Communications, Japan)
Tota Mizuno (The University of Electro-Communications, Japan)

OS22-1 Improvement of Classification Performance for Two Types of Voluntary Blinks Using a 3D CNN

Hironobu Sato[1], Shogo Matsuno[2] and Kiyohiko Abe[3]
([1]Kanto Gakuin University, Japan)
([2]University of Electro-Communications, Japan)
([3]Tokyo Denki University, Japan)

Among different types of information that can be obtained from a user's face, blinking is one of the primary actions that can be performed actively. To construct an input interface based on blinking, conscious voluntary blinking must be distinguished from naturally occurring blinking. In this study, we investigated the classification performance of a 3-dimensional convolutional neural network (CNN) model targeting voluntary and involuntary blinks. The two types of voluntary blinks adopted were blinks that subjects performed “firmly” or blinks that subjects performed “firmly for a brief time,” based on specific instructions. Furthermore, we applied techniques to improve the classification performance examined in our previous research on voluntary blink classification using a 3D CNN to verify whether these techniques remain effective even when two types of voluntary blinks are used. The improvement techniques examined in this study were automatic trimming based on the position of the center of gravity and mode value correction of the classification results. After dividing the dataset into 10 subjects for training and validation, and the remaining 5 subjects for testing, the results show the proposed method achieved an accuracy of 98.2% and an F-score of 88.1% under the conditions in which both improvement techniques were applied.

OS22-2 Individual Differences in Head Sway Variance for Drowsiness Estimation

Shunki Suzuki and Hisaya Tanaka
(Department of Information Design, Informatics Program, Kogakuin University, Japan)

This study aimed to clarify the influence of individual differences on drowsiness estimation by analyzing variations in head sway across multiple measurement sessions for each participant. A smartphone camera and MediaPipe were employed for pose estimation, and the standard deviation of head sway along the yaw, pitch, and roll axes was computed from the obtained coordinate data. The analysis demonstrated intra-participant reproducibility of head sway patterns during the drowsy state, while the predominant axis of change varied among individuals. Many participants exhibited a decrease in head sway along the yaw axis and an increase in compensatory vertical motion along the pitch axis during drowsiness. These findings highlight the necessity of designing objective indicators that account for Individual characteristics in head sway responses. Future work will perform integrated analysis combining facial expression features and head sway metrics to develop a personalized and high-accuracy drowsiness classification model.

OS22-3 Evaluation of Heart Rate and Pulse Amplitude Estimation Based on R–G Components Extracted from Facial Images

Jiawen Chen, Miku Shimizu, Tota Mizuno, Shogo Matsuno, Kazuyuki Mito and Naoaki Itakura
(The University of Electro-Communications, Japan)

This study proposes a real-time physiological estimation method based on time-domain analysis using the R–G signal extracted from facial videos. Unlike conventional FFT-based approaches that require fixed analysis windows and limit real-time performance, the proposed method directly extracts cardiac components using a fixed-band FIR band-pass filter. Heartbeat intervals are calculated from zero-crossing points, and pulse amplitude is obtained from the difference between successive peaks and troughs. Experiments were conducted on nine healthy adults under three conditions: resting, cervical vascular compression, and post-compression rest. Facial video, photoplethysmogram (PPG), and electrocardiogram (ECG) were recorded simultaneously for validation. Results show that the proposed method achieves accurate real-time heart rate estimation, with a mean absolute error within 2 beats per minute over a 1-minute interval compared to ECG across all conditions. These results demonstrate that the R–G signal enables robust and practical non-contact heart rate monitoring without frequency-domain analysis. Future work will extend this approach to stress evaluation by analyzing pulse amplitude variability.

OS22-4 Evaluation of Health Condition Based on Facial Skin Temperature Distribution Considering Occluded Regions Using Dimensionality Reduction Methods

Mahiro Hattori, Masahito Takano, Kent Nagumo and Akio Nozawa
(Aoyama Gakuin University, Japan)

This study aims to analyze distributional structures in facial thermal images related to occlusion and health condition. Principal Component Analysis (PCA) was used as a preprocessing step, and Uniform Manifold Approximation and Projection (UMAP) was applied to obtain low-dimensional embeddings. For occlusion, health condition, and health condition under non-occluded data, all combinations of the 20-dimensional UMAP output were evaluated, and the dimensional subsets yielding the highest silhouette scores were selected. Across all conditions, the average silhouette scores were close to zero, indicating that clear cluster structures were not formed in the low-dimensional space. However, the eyeglasses class showed a relatively higher score compared with other occlusion types, suggesting that eyeglasses introduce a distinct thermal pattern that is more easily separable. In contrast, health condition did not form meaningful clusters under either full-data or non-occluded settings, likely reflecting its continuous nature and the limitations of using raw pixel values as features. These results highlight the need for feature representations that preserve spatial structure in facial thermal images. Future work will incorporate spatially informed approaches, such as Local Moran's statistics and Self-Organizing Maps (SOM), to capture thermal patterns associated with occlusion and health condition more effectively.

OS22-5 Simultaneous Measurement of Muscle Blood Flow and EMG under Sub-threshold EMS

Yu Ikarashi, Kent Nagumo and Akio Nozawa
(Aoyama Gakuin University, Japan)

Sub-threshold electrical muscle stimulation (EMS), in which the stimulus is perceived but no conscious muscle contraction is reported, has been proposed as a low-burden method for supporting recovery from muscle fatigue. However, whether sub-threshold EMS truly induces no contraction and how stimulation parameters modulate physiological responses remain unclear. This study aimed to (1) determine whether sub-threshold EMS elicits imperceptible micro-contractions, and (2) evaluate how pulse width affects recovery indices. Two healthy adults performed a fatigue task, after which sub-threshold EMS with three pulse widths (2, 5, 8 ms) was applied. Muscle activity and muscle blood flow were simultaneously recorded using EMG and NIRS. EMG analysis revealed small but distinct muscle activity in all conditions, demonstrating the presence of imperceptible micro-contractions. Mean Frequency (MF) increased with pulse width, whereas ΔTotalHb showed similar decreases across conditions, likely reflecting immediate post-fatigue hemodynamic suppression. These findings indicate that sub-threshold EMS is not electrically silent but induces minimal muscle activation, and that pulse width selectively affects electrophysiological recovery. This study provides foundational insight for optimizing low-burden EMS protocols for recovery.

January 22 (Thursday), 16:15-17:30

3F Meeting Room 32

GS15 Human-machine interaction and collaboration IV

Chair: Wen Liang Yeoh (Saga University, Japan)

GS15-1 Development and Evaluation of a Telepresence System Enabling Remote Participation in a Preventive Care Game

Tamaki Kobayashi, Akira Urashima, Shin Morishima and Tomoji Toriyama
(Faculty of Information Engineering, Toyama Prefectural University, Japan)

Japan is facing rapid population aging, and many older adults experience social isolation due to physical constraints that limit their ability to go outside. This isolation is linked to declines in well-being and increased risks of cognitive impairment. Although community-based preventive care events have been promoted, participation remains difficult for older adults with mobility limitations. To address this issue, this study proposes a telepresence system that enables remote participation in a preventive care game, Master of Window Wiping , which is known to promote communication among older adults.The system allows remote participants to view 360° video of the venue through a head-mounted display (HMD) and communicate with on-site participants via audio. To enhance telepresence, a PTZ camera at the venue moves in synchronization with the remote participant's head orientation, indicating their direction of attention. An experiment was conducted during a local preventive care event to compare on-site and remote participation. The results showed that the system provided a certain level of telepresence and opportunities

for social interaction, although these were lower than those of on-site participation. As future work, it is necessary to investigate effective methods for sharing gaze and body movements, as well as to reduce video and audio latency.

GS15-2 Optimal HMI for Take-Over Requests of Autonomous Driving Systems

Ririka Kimura and Ivan Tanev
(Department of Information and Computer Science, Graduate School of Science and Engineering, Doshisha University, Kyoto, Japan)

In this study we propose a safety-first guideline for the design of the human-machine interface (HMI) for take-over requests (TOR) in autonomous driving. A set of two-stage experiments confirmed that safety must supersede promptness. The optimal – multi-modal – HMI, and a warning method involving a combination of a text message, vibration, and audio notification, achieved the minimum excessive braking risk, demonstrating superior promptness and subjective balance. We conclude that integrating clear communication channels with a calibrated tactile channel provides the most effective and safe solution for manual transition under the conditions when the human driver is engaged in non-driving related tasks (NDRT).

GS15-3 Analysis of the Association between Wrist Trajectories during Human-Human Conversation and Big Five Personality Traits using Elliptical Fitting

Xuhai Li[1], Yutaka Nakamura[2] and Yuya Okadome[1]
([1]Department of Information and Computer Technology, Tokyo University of Science, Japan)
([2]Information R&D and Strategy Headquarters Riken, Japan)

Developing natural and engaging conversational robots requires incorporating spontaneous non-verbal cues, such as unintended body language. A crucial part is the unconscious wrist movement during human conversation. This study investigates the association between these spontaneous wrist trajectories and the Big Five personality traits to insight more "natural" robot behavior design. We propose a quantitative workflow to analyze the relationship between wrist movements and individual personalities. Using video-based motion extraction, we apply elliptical fitting to characterize the shape and dynamics of wrist trajectories, and analyze these feature's correlation with participants' Big Five personality scores. Experimental results reveal a positive correlation between wrist movement frequency and Conscientiousness. Furthermore, overall speed and elliptical trajectory's area are found to be closely related to Agreeableness and Extraversion. This research provides a quantitative workflow for modeling non-verbal behaviors, offering valuable insights for designing more empathetic and human-like embodied conversational agents in human-robot interaction.

GS15-4 Discrimination of Stress Coping Patterns Based on Spatial Autocorrelation Features of Near-Infrared Facial Image

Eitaro Gono, Shonosuke Ohyama, Kent Nagumo and Akio Nozawa
(Aoyama Gakuin University, Japan)

Assessing stress coping responses classified into active and passive coping is crucial for understanding human physiological states. Conventional methods relying on contact-based hemodynamic indices are limited by measurement delays and physical constraints. This study proposes a non-contact method to discriminate stress coping patterns using near-infrared facial images(NIRFI), which reflect skin blood flow dynamics. We focus on spatial autocorrelation to extract hemodynamic features from these images. Specifically, we applied the Local Moran's I (LMI) statistic to NIRFI obtained during active (gaming) and passive (cold pressor test) coping tasks. A Support Vector Machine (SVM) was trained using the LMI images as features. The results yielded a classification accuracy of approximately 73%. Furthermore, visualizing the SVM weight matrices revealed distinct facial regions associated with each coping style: the inner canthus, nasal tip, and upper lip for active coping, and the lower lip, chin, and nasal alae for passive coping. These findings demonstrate that calculating spatial autocorrelation features from NIRFI is an effective approach for the non-contact discrimination of stress coping patterns.

GS15-5 A Study on the Potential for Drawing Robots to Induce Latent Human Movement

Asuka Yoshida and Katsuyoshi Tsujita
(Tottori University, Japan)

This study investigated two aspects using a relative drawing robot: how consistency in drawing actions is maintained, and how conscious drawing movements change in response to external stimuli. The experimental results suggested that unconscious drawing movements are not merely passive reactions but are shaped by visual information and internal processes aimed at maintaining behavioral consistency. Leveraging human compensatory movements, as demonstrated in this study, holds promise for reducing operational burden and achieving natural operation. Furthermore, in sports and skill-training devices, it is expected to elicit the individual's movement characteristics and promote actions tailored to the person.

January 22 (Thursday), 16:15-17:30

B1 Meeting Room 1

GS21 Machine learning IV

Chair: Michiharu Maeda (Fukuoka Institute of Technology, Japan)

GS21-1 Generation of Pseudo Chest CT and Ground Truth Images for Improving Opacity Segmentation Ability in Diffuse Lung Diseases

Koya Inoue[1], Shingo Mabu[1], Satoru Ikebe[1] and Shoji Kido[2]
([1]Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Japan)
([2]Institute for Radiation Science/Graduate School of Medicine, The University of Osaka, Japan)

Medical image segmentation plays a crucial role in improving diagnostic accuracy and efficiency. With advances in deep learning technology, significant improvements in segmentation performance are highly anticipated. However, its performance depends on large amounts of annotated training data, which are costly and time-consuming due to the necessity of pixel-level annotation by expert radiologists. Moreover, conventional augmentation is unsuitable since it can distort anatomical structures or alter lesion appearance. To address this problem, we propose a method for generating pseudo chest CT images and their corresponding ground truth labels for training segmentation networks. In our method, we extract lesion areas from real chest CT images and apply augmentation. Next, the augmented lesions are composited onto normal chest CT images to create pseudo images. This process allows the generation of a large variety of annotated training data without additional manual labeling effort. In verification, we trained a convolutional neural network model using the generated pseudo images and evaluated its performance on five representative opacities of diffuse lung diseases. Our experiments demonstrated that incorporating these pseudo images improves segmentation accuracy compared to training only on the real chest CT images.

GS21-2 Ensemble Learning of UTH-BERT for Anatomic Classification of Sentences of PET-CT Image Findings

Ibuki Yamanaka[1], Shingo Mabu[1], Satoru Ikebe[1] and Shoji Kido[2]
([1]Graduate School of Sciences and Technology for Innovation ,Yamaguchi University,Japan, Japan)
([2]Institute for Radiation Science/Graduate School of Medicine, The University of Osaka, Japan)

The development of AI-based computer-aided diagnosis models has been actively pursued in the medical field. However, maximizing the performance of these models requires a large amount of high-quality annotated data. Currently, preparing a sufficient amount of labeled data for every organ and disease is difficult, and the annotation process by radiologists entails significant time and labor costs. Therefore, it is necessary to develop efficient model training methods and annotation systems that utilize limited labeled datasets. Since AI applications in medical data typically focus on specific body parts or disease types, a previous study proposed a model that uses UTH-BERT to classify whole-body image findings into respective body parts. Previous investigations included adding a function for normal/abnormal classification and sharing feature extractors between different tasks. As a result, while the model demonstrated high overall classification performance, a discrepancy in classification accuracy was observed between body parts with many labels and those with few labels. In this study, we propose an ensemble learning method that trains models sequentially, aiming to improve the classification performance for body parts with limited labels. Evaluation of classification performance using AUC demonstrated that the proposed method achieved superior performance compared to the conventional method.

GS21-3 Federated Multi-Task Learning for Diffuse Lung Disease Classification and Anomaly Detection

Minori Iwase[1], Shingo Mabu[1], Satoru Ikebe[1] and Shoji Kido[2]
([1]Graduate School of Sciences and Technology for Innovation, Yamaguchi University , Japan)
([2]Institute for Radiation Science/Graduate School of Medicine, The University of Osaka, Japan)

Various deep learning models have been proposed for the purpose of classifying opacities of diffuse lung diseases in chest CT images. While these classification models demonstrate high accuracy, their construction requires large amounts of labeled training data. However, in the medical field, it is difficult to prepare a sufficient amount of labeled data because data cannot be widely shared among medical institutions due to privacy concerns. To address this issue, federated learning has been proposed as a concept that allows for the collaborative training of models while keeping data distributed across individual institutions. Based on the idea of federated learning, we applied it to build an chest CT image encoder for feature extraction aimed at both the classification and anomaly detection in chest CT images. From the experimental results, we found that the performance surpassed that of training at a single institution and was comparable to that of training with aggregated shared data. Furthermore, assuming a scenario where each institution performs different tasks, we investigated whether sharing the encoder across institutions and tasks would improve the performance of each task, and found that anomaly detection performance improved by task sharing.

GS21-4 Proposal of Price Prediction Models for Farmed Yellowtail Using Machine Learning

Hirofumi Miyajima[1], Daiki Togawa[2], Kazuki Fukae[3], Kenichi Arai[1], Tetsuo Imai[1], Mitsuru Hattori[1], Hideyuki Takahashi[4] and Toru Kobayashi[5]
([1]Graduate School of Integrated Science and Technology, Nagasaki University, Japan)
([2]School of Information and Data Sciences, Nagasaki University, Japan)
([3]Faculty of Applied Information Technology, Nagasaki Institute of Applied Science, Japan)
([4]Faculty of Informatics, Tohoku Gakuin University, Japan)
([5]Faculty of Global Media Studies, Komazawa University, Japan)

Aquaculture is a key sector of the fisheries industry. In recent years, efforts to enhance efficiency and automate tasks using artificial intelligence (AI) have expanded across various fields, include aquaculture. In this context, it is desirable to identify periods when the price of the target farmed fish is likely to increase. A central question is how to accurately predict the prices of farmed fish. The purpose of this study is to develop models that predict the price of farmed yellowtail using machine learning, a type of AI technology. The model inputs consist of historical price and harvest volume data for farmed yellowtail, as these factors are considered relevant to price fluctuations. The model output is the predicted price. Although various AI models exist, this study employs a linear function, neural network, support vector regression, and gradient boosting regression tree. Numerical experiments were conducted using actual farmed yellowtail data, and the effectiveness of the developed models was evaluated.

GS21-5 A Deep Learning-Based Estimation Method for Core Body Temperature from Skin Temperature Considering Ambient Environment

Tasuku Hanato, Shin Morishima, Akira Urashima and Tomoji Toriyama
(Faculty of Information Engineering, Toyama Prefectural University, Japan)

Due to global population aging, the resulting caregiver shortage makes it difficult to effectively conduct care monitoring for early anomaly detection in nursing facilities. Core body temperature is an effective indicator for early anomaly detection; however, because it requires physical contact, it is not suitable for continuous monitoring. Therefore, this paper proposes core body temperature estimation method from non-contactly acquired skin temperature and ambient environmental information. Because the way skin temperature changes in response to variations in ambient conditions and core body temperature differs significantly among subjects, we propose subject-specific LSTM models with five features: skin temperature, ambient temperature, humidity, atmospheric pressure, and time of day. In the evaluation, data spanning 58 hours were collected from two subjects, including both normal states and periods of stimulant ingestion designed to induce fluctuations. The proposed method was compared with models trained on multi-subject and cross-subject data. The mean absolute error (MAE) was 0.24 °C for the proposed method, compared with 0.45 °C for the multi-subject-trained model and 0.52 °C for the cross-subject–trained model. These results demonstrate that the proposed method successfully improved accuracy.

January 22 (Thursday), 16:15-17:45

B1 Meeting Room 2

GS27 Mobile robots III

Chair: Kunihiko Nakazono (University of the Ryukyus, Japan)

GS27-1 Indoor SLAM Performance Analysis of a Disaster-Rescue Tracked Robot Equipped with Dual Livox MID360 LiDAR

Juewei Yu, Jehun Seo and Yoshiaki Yamazaki
(Meisei University Graduate School of Sclence and Engineering, Japan)

This study aims to leverage the compact size and wide-area scanning capabilities of the LIVOX Mid-360 LiDAR for indoor laser mapping following disasters. To ensure scanning efficiency, a radar is mounted at both the front and rear of the robot. Through experiments in controlled environments, the impact of different geometric structures on SLAM accuracy is analyzed.

GS27-2 3D-LiDAR Point Cloud Data Classification Using Deep Learning

Abhijeet Ravankar[1], Ankit A. Ravankar[2] and Arpit Rawankar[3]
([1]Kitami Institute of Technology, Hokkaido, Japan)
([2]Tohoku University, Japan)
([3]Thakur College of Engineering and Technology, India)

Navigation is an essential component of mobile robots. Typically, robots are equipped with visual sensors to identify objects in the vicinity for obstacle avoidance. However, visual sensors have several limitations like performance degradation in low illumination or depth perception. 3D Lidars with many layers (32, 64 or 128) which generate dense point cloud data (PCD), and have shown good performance in object detection to overcome such limitations. However, 3D Lidars with fewer layers generate sparse data which are difficult for object detection. In this papers, sparse point cloud data is used for object identification using deep learning. Sparse PCD collected for different obstacle configurations (person, multiple person, bicycle) is augmented and passed into an neural network for identification. We achieved good results with sparse PCD for different obstacle configurations. Data was collected using actual sensor and classification results are discussed.

GS27-3 Introduction to position estimation Algorithm for Autonomous Vehicles Using Sensor-Fusion and AI Learning Models

Kyungho Ko and Hyunhwan Jeong
(Korea University, South Korea)

This paper proposes a sensor fusion-based position interpolation method employing artificial intelligence (AI) to minimize dead reckoning (DR) errors in Global Navigation Satellite Systems (GNSS)-degraded environments. Unlike conventional Kalman filter-based approaches, the proposed method utilizes Deep Neural Networks (DNN) and Long Short-Term Memory (LSTM) models to learn motion characteristics from onboard sensors, including accelerometers and gyroscopes. By predicting velocity at the next time step based on current and historical data, the algorithm computes displacement to accurately interpolate position. This approach effectively mitigates cumulative errors and environmental dependencies inherent in traditional methods. The validity of the proposed algorithm was verified through experiments using a multi-sensor integrated mobile platform.

GS27-4 Proposal for an Efficient Underwater Camera Calibration System Using AprilTag

Luka Fujimori, Makoto Morito and Junichiro Tahara
(Tokyo University of Marine Science and Technology, Japan)

For image analysis such as AprilTag detection underwater, camera calibration data are essential. In this study, a calibration board with AprilTags of 40 [mm] side length and different IDs arranged in 4 rows and 5 columns (hereafter referred to as a 4 × 5 array) is used. In previous study, camera calibration was performed using 27 still images in which all AprilTgs on the calibration board were detected at various positions and orientations relative to the camera. However, capturing enough images that satisfy these conditions underwater requires considerable working time. To improve calibration efficiency, this study proposes a method that utilizes image regions where any consecutive 3 × 4 array of AprilTags on the calibration board is detected. The results of underwater camera calibration obtained using the previous method and the proposed method are evaluated. Furthermore, using the camera parameters obtained through the proposed method, the AprilTags placed on the ROV are detected and the three-dimensional (3D) position and orientation of the ROV are obtained. Camera calibration and obtaining 3D information by AprilTag detection were performed in MATLAB, and analysis of the resulting data demonstrated favorable results.

GS27-5 A Gesture Recognition System for Mobile Robots with Operator Identification by Collating Camera and IMU Data

Haruaki Goto and Hirokazu Matsui
(Department of Mechanical Engineering, University of Mie, Japan)

The objective of this study is to construct a mobile robot operation system that accurately identifies a specific operator in a dynamic environment with multiple people. Addressing the challenge that individual identification is difficult with conven tional vision-based methods, this study proposes a “Visual-IMU Sensor Fusion Score” that fuses Vision-based Pose Estimation and a wearable device (IMU). This method robustly identifies only the authorized operator by integrating and evaluating three elements: “Time Synchronization (ST),” “Motion Energy (SE),” and “Posture Matching (SP).”

GS27-6 Lightweight Topological Clustering for Real-Time 3D LiDAR Mapping

Ryosuke Ofuchi, Yuichiro Toda and Takayuki Matsuno
(Graduate School of Environmental, Life, Natural Science and Technology, Okayama University, Japan)

Topological maps provide a compact representation for online 3D LiDAR mapping, yet scaling downstream recognition over long-term, incrementally accumulated maps remains computationally expensive as the mapped point cloud grows. This paper introduces a lightweight preprocessing stage that suppresses dominant planar structures and extracts recognition-relevant non-planar candidates while preserving real-time operation. We propose an online framework that integrates background masking and object-candidate clustering into topological map building based on Adaptive Resonance Theory (ART). The method leverages node-level geometric statistics estimated from assigned points and graph connectivity to separate large planar components that dominate many scenes, such as roads, floors, and walls, from residual non-planar structures. To keep the additional cost bounded under incremental updates, geometric decisions are recomputed only for locally affected graph neighborhoods, and the resulting subgraphs directly yield background-suppressed candidate sets and coarse clusters for downstream recognition. Experiments on the SemanticKITTI dataset demonstrate that the proposed approach maintains real-time incremental mapping while substantially reducing the point cloud passed to recognition and retaining non-planar targets.

January 22 (Thursday), 16:15-17:45

B1 Meeting Room 3

GS29 Multi-agent systems

Chair: Koichi Moriyama (Nagoya Institute of Technology, Japan)

GS29-1 Multi-Robot Task Planning for Multi-Object Retrieval Tasks with Distributed On-Site Knowledge via Large Language Models

Kento Murata[1], Shoichi Hasegawa[1], Tomochika Ishikawa[1], Yoshinobu Hagiwara[2], Akira Taniguchi[3], Lotfi El Hafi[4] and Tadahiro Taniguchi[5]
([1]Graduate School of Information Science and Engineering, Ritsumeikan University, Japan)
([2]Faculty of Science and Engineering, Soka Univ,
Research Organization of Science and Technology, Ritsumeikan Univ, Japan)
([3]College of Information Science and Engineering, Ritsumeikan University, Japan)
([4]Research Organization of Science and Technology, Ritsumeikan University, Japan)
([5]Research Organization of Science and Technology, Ritsumeikan Univ,
Graduate School of Informatics, Kyoto Univ, Japan)

It is crucial to efficiently execute instructions such as "Find an apple and a banana." or "Get ready for a field trip," which require searching for multiple objects or understanding context-dependent commands. This study addresses the problem of assigning subtasks to multiple robots when each robot possesses different situational on-site knowledge, specifically spatial concepts learned from the area designated to it by the user. We propose a task-planning framework that leverages large language models (LLMs) and spatial concepts to decompose natural-language instructions into subtasks and allocate them to multiple robots. We design a few-shot prompting strategy that enables LLMs to infer required objects from ambiguous commands and decompose them into executable subtasks. In experiments, the proposed method achieved 47/50 successful assignments, outperforming random assignment (28/50) and commonsense-based assignment (26/50). We also conducted qualitative evaluations using two mobile manipulators, demonstrating that the framework can handle ambiguous instructions, including ad hoc categories such as "Get ready for a field trip," through task decomposition, assignment, sequential planning, and execution. For reproducibility, we release the full set of prompts on the project website.

GS29-2 LLM-Guided Decentralized Exploration with Self-Organizing Robot Teams

Hiroaki Kawashima[1], Shun Ikejima[1], Takeshi Takai[2], Mikita Miyaguchi[2] and Yasuharu Kunii[3]
([1]University of Hyogo, Japan)
([2]Takenaka Corporation, Japan)
([3]Chuo University, Japan)

When individual robots have limited sensing capabilities or insufficient fault tolerance, it becomes necessary for multiple robots to form teams during exploration, thereby increasing the collective observation range and reliability. Traditionally, swarm formation has often been managed by a central controller; however, from the perspectives of robustness and flexibility, it is preferable for the swarm to operate autonomously even in the absence of centralized control. In addition, the determination of exploration targets for each team is crucial for efficient exploration in such multi-team exploration scenarios. This study therefore proposes an exploration method that combines (1) an algorithm for self-organization, enabling the autonomous and dynamic formation of multiple teams, and (2) an algorithm that allows each team to autonomously determine its next exploration target (destination). In particular, for (2), this study explores a novel strategy based on large language models (LLMs), while classical frontier-based methods and deep reinforcement learning approaches have been widely studied. The effectiveness of the proposed method was validated through simulations involving tens to hundreds of robots.

GS29-3 Shuttle operation decision support system using surrogate model with graph-theoretic features

Petr Smid[1] and Itsuki Noda[2]
([1]Charles University, Prague, Czechia)
([2]Hokkaido University, Japan)

In this work we address the Taxi–Shuttle Allocation Problem (TSAP), where the goal is to allocate a fixed fleet of drivers between on-demand taxis and fixed-route shuttles in order to minimize passenger cancellations under a given demand snapshot. Driver allocation directly affects how many passengers can be served and how often requests are canceled, but in practice exhaustively testing every possible split with a detailed simulator is too slow for planning. To avoid repeated simulations, we propose a learned surrogate based on ensemble tree models that predicts the cancellation ratio for candidate TSAP splits using graph-theoretic features that encode spatial coverage, connectivity, and demand–supply balance. Training scenarios are generated from the Kyoto GPS mobility dataset and mapped onto the grid to retain the empirical demand distribution while simplifying the urban layout. Experiments show that adding graph-theoretic features reduces prediction error by approximately 16% relative to non-graph surrogates. The surrogate also outperforms domain-informed heuristics, while requiring only a fraction of the simulation computation time. These results support graph-based surrogate modeling as a scalable alternative to simulation-driven solutions for TSAP in systems that combine taxis and shuttles.

GS29-4 Comparison of Preference Weight Input Methods in Multi-Objective Multi-Agent Reinforcement Learning

Yoshiki Nogata[1], Tomohiro Harada[1] and Fumito Uwano[2]
([1]Graduate School of Science and Engineering, Saitama University, Japan)
([2]Faculty of Environmental, Life, Natural Science and Technology, Okayama Univerisity, Japan)

This study analyzes the behavior of MO-MIX, a representative method for multi-objective multi-agent reinforcement learning (MOMARL), under different input strategies for assigning preference weights to objectives during training. Specifically, in addition to random preference-weight sampling used in the previous study, we compare order-based weight-input strategies, including ascending, descending, and variants that reverse the order at the maximum and minimum weights. To investigate the impact of weight input strategies on the learning performance on MO-MIX, we conduct experiments on a two-objective, three-agent benchmark problem. The experimental results show that while certain ordered strategies achieve slightly higher performance on specific metrics, no weight-input strategy consistently outperforms the random baseline across all criteria.

GS29-5 A Control Strategy for Monitoring Unknown Flood Regions by Multiple UAVs

Jie Song[1], Yang Bai[2], Mikhail Svinin[3], Evgeni Magid[4] and Naoki Wakamiya[1]
([1]Graduate School of Information Science and Technology, The University of Osaka, Osaka, Japan)
([2]Graduate School of Advanced Science and Engineering, Hiroshima University, Hiroshima, Japan)
([3]Graduate School of Information Science and Engineering, Ritsumeikan University, Osaka, Japan)
([4]Institute of Information Technology and Intelligent Systems,
Kazan Federal University and HSE University, Moscow, Russia)

This study introduces an innovative control approach for deploying multiple unmanned aerial vehicles (UAVs) to monitor an unknown flood region. The proposed strategy is designed to optimally distribute UAVs across the flood-affected area while cooperatively estimating the extent of inundation. To achieve this, an adaptive coverage controller is developed based on Centroidal Voronoi Tessellation (CVT), incorporating a novel mechanism for dynamically updating the density function. Within this framework, the density function serves as an evolving representation of the estimated inundation areas, allowing UAVs to adjust their positions adaptively in response to real-time environmental changes. The effectiveness of the proposed control strategy is validated through simulations conducted in the ROS/Gazebo environment, demonstrating its capability to enhance the accuracy of flood monitoring and improve the spatial distribution of UAVs.

GS29-6 BST-ID–Guided Broadcast: Prefix-Scoped Diffusion for LPWAN Mesh Networks

Yuki Nagasawa and Yuichi Yaguchi
(University of Aizu, Japan)
(University of Aizu, Japan)

Epidemic-style broadcast is a practical choice for LPWAN mesh networks because it works under intermittent connectivity and unknown, dynamic topology. However, naive flooding quickly becomes inefficient: duplicate relays increase airtime usage and collisions, especially as node density grows. This paper presents a BST-ID–guided broadcast method that enables lightweight, scope-aware diffusion for LoRa-like multi-hop LPWAN meshes. Each packet carries a Binary Spatio-Temporal Identifier (BST-ID) representing the intended destination context. Upon reception, intermediate nodes make a local forwarding decision using three simple checks: duplicate suppression (first reception only), hop limitation (TTL), and a BST-ID–guided progress rule that suppresses relays that do not move closer to the target in terms of decoded, approximate distance. We evaluate the method using a discrete-event simulator with slotted ALOHA medium access and a simplified destructive-collision model. Results show that BST-ID guidance reduces diffusion footprint, overall transmissions, and collision pressure, while maintaining effective target reachability in sufficiently dense deployments. The study highlights diffusion control as a key design dimension for scalable LPWAN dissemination without routing tables or centralized coordination.

January 22 (Thursday), 16:15-17:45

B1 Meeting Room 4

GS33 Robotic Mechanism

Chair: Shuhei Ikemoto (Kyushu Institute of Technology, Japan)

GS33-1 Effect of Radial Clearance on the Reaction Moment in Cylindrical Shaft–Bushing Pairs

Hikaru Koyama, Tatsuro Terakawa, Ryohei Michikawa and Masaharu Komori
(Department of Mechanical Engineering and Science, Kyoto University, Japan)

This research examines the applicability of Hertzian elliptical contact theory to a cylindrical shaft-bushing pair with radial clearance, a configuration frequently encountered in robotic sliding mechanisms. A Hertzian formulation is constructed to compute the reaction moment from the normal contact solution using local principal curvature radii near the bushing fillet. The elastic-contact assumption is checked by evaluating the maximum pressure, and the small-contact condition is quantified by the dimensionless ratios between the semi-axes and the relative radii of curvature obtained from the Hertzian solution. Finite element method (FEM) contact analyses are also performed under identical geometry and material conditions. While the Hertzian model predicts strong sensitivity of the reaction moment to the fillet radius and only weak dependence on clearance, FEM results exhibit pronounced clearance dependence. A comparison of contact areas shows that FEM predicts a contact region substantially larger than the Hertzian elliptical patch, indicating that the local elliptical-patch assumption may break down even when elastic and small-contact conditions are satisfied. These findings clarify practical limits of Hertzian theory for clearance-dependent prediction of reaction moments in

GS33-2 Development of a Desk Organization Robot and Its End-Effector

Natsuki Muta[1], Hideaki Itoh[1], Hisao Fukumoto[1] and Hiroshi Wakuya[2]
([1]Electrical and Electronic Engineering Course, Graduate School of Science and Engineering, Saga University, Japan)
([2]Faculty of Education, Saga University, Japan)

Cluttered work environments have been shown to negatively impact mental health and work efficiency. Accordingly, this study aims to develop a robotic system capable of autonomously organizing objects on a desk. Conventional grasping methods often require object-specific joint angle tuning, while approaches based on pre-trained models tend to exhibit reduced accuracy when handling previously unseen objects. In the proposed system, an RGB-D camera is used to capture a reference state of the desk, and object displacement is identified through difference detection relative to this baseline. Object recognition is performed using the You Only Look Once (YOLO) algorithm, which provides real-time performance and strong generalization capabilities. The detected object coordinates are then transformed into the robotic arm's coordinate system, enabling accurate grasping and repositioning of objects. Furthermore, a two-finger end-effector incorporating compliant rubber material is developed. This design allows stable grasping of objects with diverse shapes without the need for force sensors or complex control algorithms.

GS33-3 The Effect of Lug Alignment Angle on Driving Performance and Evaluation of Lug Installation Conditions for Driving Performance Improvement ~ Two-track Crawler Running Experiment and PIV Analysis~

Kazuki Maeda[1] and Katsuyoshi Tsujita[2]
([1]Graduate School of Tottori University, Japan)
([2]Tottori University, Japan)

The surface of the Moon and other celestial bodies is soft ground, making it prone to a phenomenon called "stuck" where the wheels spin freely, and the rover becomes unable to move when it drives over it. Especially near craters and at the poles, there are areas with rugged terrain and steep slopes. Improving mobility is essential to traverse such terrain. However, the optimal conditions for improving mobility have not been clearly established. This paper focuses on the installation conditions of the lugs. Among these conditions, we concentrate on the lug alignment angle. Single-track crawler experiments have shown that driving performance varies with the lug alignment angle. In this study, we constructed a two-track crawler and conducted driving experiments to validate the findings of previous research. Additionally, experiments conducted. To investigate the impact of lugs on sand during travel, PIV analysis is employed. PIV analysis visualizes the fluidic behavior of sand during movement. Based on the experimental results from the two-wheeled crawler on level and sloped terrain, along with the PIV analysis results, the relationship between travel performance and lugs is examined, providing a basis for future quantitative evaluation.

GS33-4 Development of a Compact Scanning System for LiDAR Applications Using an Miniature Ultrasonic Motor and Rotary Encoder

Tsukiho Fujitani[1], Hiroaki Fukushima[1] and Shunsuke Izuhara[2]
([1]Kyoto University of Advanced Science, Japan)
([2]Okayama University, Japan)

Compact and high-precision LiDAR systems are increasingly required for autonomous robots and medical imaging devices, where the available installation volume and power budget are severely limited. This study proposes a miniaturized LiDAR-type scanning architecture driven by a 2.0 mm ultrasonic motor whose stator has a simple configuration composed of a single metal plate and a single piezoelectric element, enabling straightforward integration into compact sensing units. The motor is operated near its flexural resonance to generate a traveling wave on the stator surface, which converts to rotor rotation through frictional contact. A rotary encoder is incorporated as an input device for command generation, and a PID-based duty-ratio control scheme is implemented to achieve precise angle positioning with an angular resolution of 0.27°. In addition, the proposed framework supports synchronized operation with external sensors, such as a camera and a Time-of-Flight sensor, by using a miniature mirror integrated on the rotor to realize optical scanning without bulky external scanning mirrors. Experimental demonstrations confirm stable and repeatable positioning and synchronized motion, indicating that the proposed system provides a promising platform for compact scanning modules in endoscopic imaging and small-scale robotic sensing.

GS33-5 Flexible suction device for maintaining position and orientation of underwater robots

Yuya Takayama and Norimitsu Sakagami
(Ryukoku University, Japan)

Regular inspections of underwater structures are crucial for preventing deterioration and damage. Traditionally, these inspections are performed by divers. However, underwater robots are increasingly replacing divers. Free-floating underwater robots are vulnerable to wave impacts and struggle to maintain their position and posture in underwater environments. To address this challenge, some underwater robots are equipped with position-keeping devices to maintain their position. We designed and implemented a suction device that maintained mechanical contact with structural surfaces. Previously, the suction device was specifically designed to attach to flat surfaces. This study aims to design and develop a new suction device capable of generating a strong suction force on different surfaces. The contact part of the proposed device was made of soft material. The suction performance was evaluated through several experiments using both rigid and flexible suction devices. The results confirmed that the flexible suction device generated a higher suction force than the rigid device.

GS33-6 Design and Development of Mode-Switching Linear–Rotary Actuation Mechanism for Enhanced Versatility of Closed-Chain Bipedal Legs

Mohnish Ganjam Chandramohan, Khalid Meitani, Gajitha Nanayakkara and Sajid Nisar
(Kyoto University of Advanced Science, Japan)

Humanoid bipedal legs are commonly realized as open-chain mechanisms because they offer high trajectory versatility, but this typically comes at the cost of lower rigidity and higher knee loading. Closed-chain legs provide improved stiffness, stability, and energy efficiency, yet their closed-loop constraints limit versatility. To bridge this gap, this paper presents a mode-switching linear–rotary actuation mechanism that enables an adjustable link in a five-bar closed-chain leg to emulate trajectories comparable to a serial 2R leg without changing hardware. The proposed mechanism uses a single DC motor as the primary power source and a servo-driven selector to alternately (i) couple the motor to the output shaft for rotary actuation or (ii) engage a bevel-gear–lead-screw transmission to generate linear extension/contraction in an orthogonal direction. A 3D-printed prototype is fabricated and experimentally evaluated, demonstrating reliable switching between modes within 0.344–0.65 s and linear motion at 13.33 mm/s (full extension in 11.25 s under the tested conditions). These results validate the feasibility of compact, mode-selectable linear–rotary actuation as a practical route to increase closed-chain leg versatility while retaining their inherent stability.

January 23 (Friday), 9:00-10:45

3F Meeting Room 32

OS26 ISBC: TONAL 2026

Organizers: Georgii Karelin (Okinawa Institute of Science and Technology, Japan)
Milan Rybar (Okinawa Institute of Science and Technology, Japan)
Moritz Kriegleder (Okinawa Institute of Science and Technology, Japan)
Luna Wang (Okinawa Institute of Science and Technology, Japan)

Invited Talk 2 A novel formalization of the mind-body relation

Tom Froese (Okinawa Institute of Science and Technology, Japan)

See page 16

Invited Talk 3 When Decisions Matter: Entropy Dynamics in EEG as a Marker of Deliberate Choice

Milan Rybar (Okinawa Institute of Science and Technology, Japan)

See page 17

OS26-1 Sequential decision-making under uncertainty in yeast

Haruka Terauchi[1,2], Luna Wang[1], Tom Froese[1]
([1]Embodied Cognitive Science Unit,
Okinawa Institute of Science and Technology Graduate University, Japan)
([2]Research Institute, Suntory Global Innovation Center Ltd., Japan)

Like all organisms, the single-celled budding yeast aims to optimise its behaviours for survival. One key choice is when to begin reproduction during the cell cycle. Drawing from decision theory and the fact that histone mRNA levels (marking that start of the S-phase) peak at about 37% of the cell cycle, we hypothesize that yeast cells are implicitly implementing the famous 37% rule – an optimal strategy for sequential decision-making – to time the start of cell division. We propose that this timing could be governed by stochastic gene expression, whereby the rate of fluctuations serves as an adaptive “internal clock”. Here, we report that the rate of transitions in gene expression in budding yeast is near optimal for the timing of the start of the S-phase, as modelled by a Poisson process model with lambda e. Our findings suggest a novel link between gene expression fluctuations and the regulation of cell reproduction.

OS26-2 Noise as a resource in quantum biology

Moritz Kriegleder[1,2], Andrea Loettgers[1] and Tarja Knuuttila[1]
([1]Department of Philosophy, University of Vienna, Austria)
([2]Embodied Cognitive Science Unit, Okinawa Institute of Science and Technology, Japan)

Since the inception of quantum theory, pioneers such as Schrödinger, Bohr, and Delbrück have debated the implications of quantum physics for molecular biology. Today, this intersection has expanded into an interdisciplinary field, applying quantum mechanics to the chemical foundations of life, functional biological mechanisms, genetic evolution, and high-precision macromolecular experiments. In this paper, we discuss how contemporary quantum models of photosynthesis have fundamentally transformed our understanding of entanglement. The complex, dynamic environments of biological organisms demonstrate that noise can be reconceptualized as a functional resource for sustaining entanglement. This stands in contrast to the design philosophy of quantum computers, which are engineered to be isolated from environmental interactions to the greatest extent possible. Consequently, the translation of physical models into the biological domain does more than clarify processes like photosynthesis; it forces a reconsideration of the role of noise within physics. Rather than merely resisting decoherence, biological systems appear to have evolved to utilize noise as a mechanism for self-organization within unstable and unpredictable environments. We explore methodologies to move beyond the isolated, Markovian models of traditional quantum physics to incorporate essential biological characteristics such as historicity, agency, and adaptivity.

OS26-3 Philosophical and Practical Implications of Non-Determinism in Some Computations

Georgii Karelin
(Embodied Cognitive Science Unit, Okinawa Institute of Science and Technology, Japan)

Idealized and perfect machines serve as a helpful metaphor for understanding and building practical artificial tools. However, it is an impossible construction, in the same way that there are no material points, perfectly rigid bodies, or ideal gases, and other similar concepts from physics. Those ideas are still helpful, not just despite, but precisely because they are approximate and good enough to provide helpful compressed descriptions of reality in most cases, most of the time. Both mainstream pro-AI (Inflationary, Technocratic) and anti-AI (Deflationary, Humanistic) positions essentially treat errors, mistakes, noise, and variability in AI as something suboptimal and undesired. For the first group, mistakes should be eliminated; for the second group, the presence of errors provides definitive proof that artificial intelligence or artificial consciousness is impossible. This paper carefully re-analyzes the common argument that artificial intelligence is not possible in computers because computers are formal, closed, deterministic, fully observable, and perfectly predictable systems. This paper fully agree with the requirements of openness and indetermination, but argue that real machines are not identical with their ideal Platonic descriptions. It is well known that LLMs often make mistakes, ranging from minor to catastrophic, and that some powerful realizations are non-deterministic, producing variable but somewhat coherent answers. One less-known example of indeterminism already exists in present-day computational systems: the non-associativity of floating-point arithmetic in large-scale parallel computations, where slight initial variations accumulate and are amplified by autoregressive GPT architectures. Following Alan Turing and his lesser-known ideas, it might be proposed that real intelligence cannot avoid mistakes; it must manage, recover, and learn from them

January 23 (Friday), 9:00-10:15

3F Meeting Room 33

OS1 AROB: Advances in Autonomous and Cooperative Drone Systems

Organizers: Yuichiro Sueoka (The University of Osaka, Japan)
Yuki Minami (University of Hyogo, Japan)
Kazuma Sekiguchi (Tokyo City University, Japan)

OS1-1 An LLM-based Architecture for Bidirectional Human-Drone Swarm Collaboration

Simon Kah Seng Ng[1], Xing Ying Low[1], Kazuki Ito[2] and Yuichiro Sueoka[2]
([1]The University of Osaka, Japan)
([2]Graduate School of Engineering, The University of Osaka, Japan)

This paper presents a decentralized multi-UAV architecture designed for cooperative tasks such as ecological monitoring and logistical transport, enabling dynamic team organization and reorganization under uncertain and rapidly changing mission conditions. While existing swarm coordination methods rely on fixed structures or pre-defined task models, our approach integrates Large Language Models (LLMs) to introduce high-level reasoning and intent-aware adaptability into the swarm. Each drone employs a task-parsing (planning) LLM to interpret operator-issued natural-language commands into structured swarm actions, while a secondary consultant LLM monitors system states, detects potential failures, and proposes real-time adjustments or task replanning. Combined with a ROS 2-based coordinator-cooperator negotiation protocol, this dual-LLM design allows UAVs to self-organize according to mission context, maintain robustness against anomalies, and provide interpretable feedback to human operators. This design not only allows operator-swarm communication but also generates valuable feedback from real-world conditions in the field, which enhances decision support by introducing reasoning and contextual interpretation beyond conventional algorithmic methods. Experimental demonstrations using Tello EDU drones validate the feasibility and adaptability of the proposed architecture.

OS1-2 Dual-IMU Relative State Estimation for Tethered UAV Payloads using Kinematic Constraints

Arif Fahimey, Kazuma Sekiguchi and Kenichiro Nonaka
(Tokyo City University, Japan)

Most existing research on UAVs with suspended payloads relies on simplified point-mass models, often neglecting the critical 3-DOF payload orientation. Furthermore, conventional single-IMU estimation methods are model-dependent and generally cannot observe the payload's relative state without the aid of external vision systems. This work addresses this gap by proposing a dual-IMU sensor fusion framework that instruments both the UAV and the suspended payload. This architecture transforms the problem from indirect model-based inference to robust, kinematically constrained relative-state estimation. A tailored Error-State Kalman Filter (ESKF) fuses data from distributed inertial sensors to estimate the payload's 3-DOF orientation and relative position vector. A central feature of this framework is the incorporation of the tether length as a hard kinematic constraint within the measurement update. This constraint ensures the observability of the relative position—mitigating the dead-reckoning drift inherent in low-cost IMUs—while maintaining computational efficiency. The proposed system is validated on a physical platform, demonstrating accurate tracking of swing dynamics in preliminary static-base experiments. These results verify that the dual-IMU configuration provides superior observability compared to single-sensor approaches, establishing a validated baseline for future dynamic flight control.

OS1-3 Real-Time Pedestrian Avoidance for a Small Quadrotor Using YOLOv11 and Depth Anything V2

Shinnosuke Matsuyama, Kimiko Motonaka and Seiji Miyoshi
(Kansai University, Japan)

Safe navigation with moving obstacles remains a challenging problem for small quadrotors equipped with limited onboard sensors. In this paper, we present a vision-based method for avoiding moving obstacles using only monocular images acquired by a camera mounted on a quadrotor. While previous research often failed to avoid obstacles approaching diagonally due to inaccurate depth estimation, the proposed system achieves real-time estimation of the 3D positions of moving obstacles in real time by combining YOLO-based object detection with monocular depth estimation using Depth Anything V2. Based on these estimated positions, the system calculates the predicted trajectories of the moving obstacles up to several steps ahead. An appropriate avoidance maneuver is then determined by evaluating the spatial relationship between these predicted trajectories and a predefined avoidance zone. The effectiveness of the proposed method was validated through indoor flight experiments using a small quadrotor, Tello. The results confirm that the system successfully detects and avoids moving obstacles, specifically pedestrians approaching from both front and diagonal directions. This approach enables real-time obstacle avoidance in indoor environments using minimal hardware resources.

OS1-4 Simulation Study on 3D Reconstruction from Drone Imagery with Target Tracking

Hajime Tazawa, Ichiro Maruta and Kenji Fujimoto
(Kyoto University, Japan)

The use of drone imagery for 3D reconstruction has attracted growing attention in fields such as disaster response, agriculture, and infrastructure monitoring. However, most existing reconstruction methods are designed for static scenes and often fail in dynamic environments containing moving objects or motion blur. Although recent dynamic reconstruction frameworks have shown encouraging results, they typically require dense image sequences or strict temporal synchronization, which are difficult to achieve in outdoor conditions. In this study, we explore an alternative approach that exploits the mobility of drones by actively controlling their motion and camera orientation during image acquisition. We focus on View-Track imaging strategies, in which the camera continuously tracks a moving target to suppress motion blur on the subject itself, effectively confining blur to the background similar to panning photography. Simulation experiments were conducted in NVIDIA Omniverse under multiple flight modes and blur conditions. The captured images were reconstructed using DUSt3R, and the relationship between blur distribution and reconstruction quality was qualitatively analyzed. The results demonstrate that selectively suppressing blur on the target region significantly influences reconstruction performance, indicating that active target tracking is crucial for reconstructing dynamic scenes and creating practical digital twins.

OS1-5 Neural Network Approximation and Model Quantization for MPC towards Edge Implementation

Naoki Tsubone[1], Yuki Minami[2] and Masato Ishikawa[1]
([1]Graduate School of Engineering, The University of Osaka, Japan)
([2]Graduate School of Engineering, University of Hyogo, Japan)

Model Predictive Control (MPC) is an advanced control method that determines control inputs by solving an optimization problem in real time based on a model of the controlled system and its current state. However, for edge devices such as drones, which have limited computational resources and strict battery constraints, the computational burden of MPC can be prohibitive. Therefore, control methods with higher computational speed and lower power consumption are required. To address this challenge, we explore an approach in which MPC is replaced by a neural network (NN) to reduce computation time. The NN is trained offline to learn the mapping between the current state and the control input obtained from the MPC optimization. Furthermore, we investigate the quantization of the NN weight coefficients to reduce memory usage. In this paper, we evaluate the proposed approach in terms of computation time, memory usage reduction, and control performance when the NN coefficients are quantized.

January 23 (Friday), 9:00-10:00

B1 Meeting Room 1

OS5 AROB: Bio-inspired Theories and Applications (1)

Organizers: Kunihito Yamamori (University of Miyazaki, Japan)
Kentaro Aburada (University of Miyazaki, Japan)

OS5-1 Emotional Japanese Words Vectorization by Emotion-Anchored Contrastive Learning

Daisuke Amaki[1], Shuta Yamane[1] and Kunihito Yamamori[2]
([1]Graduate School of Engineering, University of Miyazaki, Japan)
([2]Faculty of Engineering, University of Miyazaki, Japan)

Modern people experience psychological stress daily and face numerous worries. It is difficult for us to confide in others about our personal worries due to shame and fear. Someone claims that Generative AI (Artificial Intelligence) can serve as an ideal personal confidant because it is not human and can keep our personal worries confidential. However, general-purpose language models capture contextual meaning well but often fail to represent emotional features, which can lead to inappropriate responses. The objective of this research is to develop a model that can vectorize sentences based on the emotions expressed in the text. For this objective, we modify Emotion-Anchored Contrastive Learning (EACL) for Japanese. The model is trained to map sentences near their corresponding emotion anchors and away from anchors of other emotions. The proposed model is evaluated from three different viewpoints. The first one is the F1 score for emotion classification. The second is classification accuracy, which is the proportion of the sentences with the same emotion as a given sample in the nearest 10 samples. The last one visualizes the emotion-anchor distributions before and after training. The model with EACL shows a clearer cluster than the pre-trained Japanese model.

OS5-2 Automatic Drive Support with Multi-tapped VLM by Dynamic Patch Selection

Kotaro Shiiki[1], Yuma Nagatomo[2] and Kunihito Yamamori[3]
([1]Graduate School of Engineering, University of Miyazaki, Japan)
([2]Comprehensive Technology Center, University of Miyazaki, Japan)
([3]Faculty of Engineering, University of Miyazaki, Japan)

Vision-Language Models (VLMs) are powerful AI systems essential for real-time applications like Advanced Driver-Assistance Systems (ADAS) and robot control, as even minor delays can lead to critical situations. However, current VLM performance is hampered by two main inefficiencies: the Vision Transformer (ViT) generates redundant data from video frames, and the Large Language Model (LLM) must wait for ViT to complete its processing entirely. Existing acceleration techniques, such as interruption in ViT and token pruning, are insufficient. To effectively address this critical trade-off between computational speed and generated text quality, we propose a novel integration of dynamic early termination of ViT with VisionZip token compression. Our proposed method inserts a predictive head early in the ViT pipeline. This head dynamically assesses input complexity and bypasses subsequent unnecessary layers, directly transferring the optimized features to VisionZip. By merging these two strategies, we significantly

reduce the volume of input tokens, thereby lowering the computational load on both the ViT and the LLM. The core objective of this research is to establish the optimal balance between accuracy and execution time for this integrated framework, which will be rigorously evaluated using the ScienceQA benchmark.

OS5-3 Long-term Consistent Conversation of Interactive AI by Similar Sentence Retrieving on Differentiable Neural Computers

Shuta Yamane[1], Yuma Nagatomo[2] and Kunihito Yamamori[3]
([1]Graduate School of Engineering, University of Miyazaki, Japan)
([2]Comprehensive Technology Center, University of Miyazaki, Japan)
([3]Faculty of Engineering, University of Miyazaki, Japan)

Conversational AI based on Large Language Models (LLMs) has a weak point in maintaining long-term consistency and memory. Since conventional methods, such as Retrieval-Augmented Generation (RAG), rely on semantic similarity, they are ineffective in ``implicit reasoning'' scenarios in which there is a long semantic distance between queries and their corresponding evidence sentences. As a result, AI agents cannot maintain long-term consistency and memory throughout the conversation. To solve this problem, we propose a novel retrieval framework based on a Differentiable Neural Computer (DNC). Our DNC-based approach is designed to learn a retrieval algorithm from the temporal order of sentences within the conversation. The DNC utilizes a ``temporal linkage mechanism'' to track contextual relationships and state transitions over time. This mechanism enables us to identify contextually relevant sentences without relying on semantic or static similarity. We evaluate our proposed method on the ImplexConv dataset, which is specifically designed to assess implicit reasoning in long-term dialogues. To assess whether our proposed model can identify evidence sentences, we employ the F1 score in complex scenarios such as Supportive and Opposed reasoning.

OS5-4 Pixel-Level anomalous behavior detection from reconstructed images by GAN

Kazuma Masuda[1] and Kunihito Yamamori[2]
([1]Graduate School of Engineering, University of Miyazaki, Japan)
([2]Faculty of Engineering, University of Miyazaki, Japan)

Numerous surveillance cameras are deployed in modern cities to ensure public safety, producing vast amounts of video data every day. Deep learning-based video anomaly detection aims to automatically identify unusual events in such footage. Among existing approaches, reconstruction-based methods using CNN generative models (e.g., autoencoders and GANs) are widely used: models are trained on normal data to reconstruct input frames, and anomalies are detected when reconstruction errors become large. However, surveillance cameras often employ wide-angle lenses to cover large areas, and anomalous objects may appear small and distant. In this setting, spatial downsampling in CNNs can weaken fine-grained cues, causing minor anomalies to be missed and degrading detection performance. Pixel-level anomaly detection is a promising direction because it evaluates normality at each pixel and can capture small, localized abnormal regions. While pixel-level detection has been effective for still-image defect inspection, its practical benefits for video anomaly detection have not been sufficiently quantified. This study evaluates GAN-based reconstruction methods from a pixel-level perspective and analyzes when pixel-wise scoring improves sensitivity to small anomalies in surveillance videos.

January 23 (Friday), 9:00-10:30

B1 Meeting Room 2

GS11 Evolutionary computations (Genetic algorithm)

Chair: Akimasa Otsuka (Sanyo-Onoda City university, Japan)

GS11-1 Analysis of the impact of Ill-Conditioned RBF Surrogates on the Optimization Performance of Surrogate-Assisted Evolutionary Algorithms

Yuto Kano[1], Yuki Hanawa[2] and Tomohiro Harada[1]
([1]Graduate School of Science and Engineering, Saitama University, Japan)
([2]Graduate School of Systems Design, Tokyo Metropolitan University, Japan)

Surrogate-assisted evolutionary algorithms (SAEAs) are practical approaches for solving computationally expensive optimization problems. Radial basis function (RBF) is a widely used surrogate model in SAEAs. Despite its widespread use, the RBF surrogate has a potential risk: as sample points become dense during the search, the kernel matrix can become ill-consitioned, wich may lead to numerical instability of surrogate prediction. This study aims to investigate how such ill-conditioning in RBF affects the prediction accuracy of the surrogate model and the overall optimization performance of SAEAs. Specifically, we evaluate the prediction accuracy of RBF surrogates under both ill-conditioned and well-conditioned scenarios during the optimization process. Furthermore, we compare the optimization performance of SAEAs under ill-conditioned and well-conditioned scenarios. The experimental results reveal that, although the RBF surrogate showed a significant decline in prediction accuracy in ill-conditioned scenarios, this had a limited impact on search performance.

GS11-2 An Evolutionary Computation Framework to Analyze the Relationship Between Constraint Settings and Optimal Solutions in Multiple-Constraint Optimization Problems

Shunto Nabata and Tomohiro Harada
(Graduate School of Science and Engineering, Saitama University, Japan)

Many real-world optimization problems are multiple-constrained optimization problems (MCOPs). In practical settings, constraint settings are often biased by the prior knowledge or experience of designers and decision-makers, which can hinder the discovery of promising solutions outside the imposed constraints. Therefore, it is crucial to analyze how optimal solutions change as constraint settings vary. To address this issue, we propose a framework that models the relationship between the constraint settings and optimal solutions in MCOPs. Specifically, we extend our previous method, which models this relationship through a search phase using evolutionary algorithms (EAs) and a prediction phase using Gaussian process regression (GPR). In this study, we introduce an enhanced prediction-phase strategy that determines the next constraint setting by synergistically leveraging two key properties of the GPR model: predictive uncertainty and gradient of regression curve. We evaluated the proposed method using two-constraint knapsack problems under three distinct scenarios: Basic, Stepwise, and Stagnation. The experimental results show that the proposed method achieved the lowest prediction error with fewer samples in the Basic and Stepwise scenarios, whereas the uncertainty-only method provides more accurate approximations in the Stagnation scenario.

GS11-3 Automatic Generation of Challenging Optimization Benchmarks via Large Language Models and Evolutionary Computation

Yuto Araki[1], Yuhiro Ono[2] and Tomohiro Harada[1]
([1]Graduate School of Science and Engineering, Saitama University, Japan)
([2]Graduate School of Systems Design, Tokyo Metropolitan University, Japan)

Benchmark functions are essential for objectively evaluating optimization algorithms, yet most existing benchmarks are manually designed and lack the diversity and complexity of real-world problems. This study proposes an automatic benchmark generation framework that integrates large language models (LLMs) with evolutionary algorithms (EAs) to address these limitations. In the proposed method, optimization benchmark problems represented as mathematical expressions are treated as individuals in an EA, and new problems are generated using an LLM as an operator analogous to crossover and mutation. Few-shot prompting enables the LLM to produce diverse candidate expressions, while the EA guides the search toward more challenging problems. The difficulty of each generated benchmark is evaluated using a fitness function based on the statistically normalized difference between the best and second-best objective values obtained across multiple trials of a given optimization algorithm,

measuring the isolation of the global optimum. A penalty term is also introduced to prevent divergence and ensure numerical stability. Experiments generating five-dimensional single-objective problems using Differential Evolution and the Llama 3 70B model demonstrate that the proposed method produces significantly more challenging problems than standard benchmarks. The results highlight the effectiveness of the framework in generating stable, diverse, and difficult optimization benchmarks automatically.

GS11-4 Set-based Differential Evolution with Exponential and Binomial Crossovers for Discrete Optimization Problem

Yuto Kuninaga[1], Michiharu Maeda[2] and Yuta Chikuba[2]
([1]Fukuoka Institute of Technology, Japan)
([2]Digital Information Technologies Corporation, Japan)

This paper presents set-based differential evolution with exponential and binomial crossovers for discrete optimization problems. While differential evolution has been traditionally applied to continuous optimization problems, it can be adopted to discrete optimization problem by extending techniques in continuous space into discrete space by employing set-based representation. In our approach, a candidate solution is defined as a crisp set, and all arithmetic operations in mutation are redefined through novel operations. The mutation operator in our algorithm adds two different solutions selected randomly to the current solution and the new solution is constructed probabilistically. For two crossovers, the number of inherited mutant parameters follows an exponential distribution and inherited mutant parameters follows a binomial distribution in terms of their number. This study investigates exponential and binomial crossover mechanisms within the framework of set-based differential evolution. Specifically, we examine the influence of crossover rate on solution accuracy in traveling salesman problem, conducted in two phases: broad search and narrow search. In order to demonstrate the effectiveness of our algorithms, we examine numerical experiments and compare results with existing algorithms.

GS11-5 GA-based edge bundling with Bézier-curve representations

Chihiro Noda and Ryosuke Saga
(Graduate School of Informatics, Osaka Metropolitan University, Sakai, Osaka, Japan)

This paper proposes an evolutionary edge bundling method that improves network readability by optimizing edge shapes via control-point placement. Unlike conventional evolutionary approaches that evaluate polylines passing through multiple control points, the proposed method evaluates candidate solutions in the quadratic Bézier curve domain, which suppresses zigzag artifacts and produces smoother bundles. For objective computation, each Bézier curve is uniformly sampled at ten interior points and approximated as a polyline. The optimization is formulated as a multi-objective problem and solved using NSGA-II with four metrics: Mean Edge Length Difference (MELD), Mean of Occupation Area (MOA), Edge Density Distribution (EDD), and Crossing Count (CC). Experiments on the Japanese airline route network (79 nodes, 233 edges) show that the proposed method reduces zigzag patterns compared with GABEB and avoids excessive absorption of peripheral edge groups observed in FDEB. These results suggest that Bézier-based evaluation is effective for evolutionary edge bundling under multiple aesthetic criteria.

GS11-6 Two-stage Identification of Continuous-time Hammerstein Systems Using Gaussian Process Models

Tomohiro Hachino and Yuki Kawauchi
(Kagoshima University, Japan)

This paper deals with a two-stage identification of continuous-time Hammerstein systems using a Gaussian process (GP) model. The Hammerstein system is described by the cascade connection of a nonlinear static part followed by a linear dynamic part. In the first stage, the nonlinear static part represented by the GP model is estimated based on multiple sets of data of constant input and corresponding steady-state output. Then, in the second stage, the linear dynamic part is estimated by the linear least-squares method based on the approximated discrete-time estimation model derived by a digital pre-filter. Since the training of the GP model is limited to estimating the nonlinear static part only, the proposed method can be applied to systems whose linear dynamic parts have arbitrary order and can significantly reduce the computational burden.Simulation results show that the proposed method provides accurate identification models and significantly reduces the computational burden for identification.

January 23 (Friday), 9:00-10:30

B1 Meeting Room 3

GS17 Intelligent control

Chair: Koichi Kobayashi (Hokkaido University, Japan)

GS17-1 Hierarchical Reinforcement Learning–Based Active Disturbance Rejection Control for Robotic Actuators

Yucheng Meng, Ruitong Liu and Kenji Hashimoto
(Graduate School of Information, Production and Systems, Waseda university, Japan)

This paper proposes a novel hierarchical control architecture that synergizes an Extended State Observer (ESO) with Proximal Policy Optimization (PPO) to address the inherent stability-responsiveness trade-off and significant phase lag in Variable Stiffness Series Elastic Actuators (VSSEAs) caused by large-range stiffness mutations. By utilizing the ESO as a dynamic linearizer to decouple model nonlinearities and employing a stiffness-aware PPO policy for adaptive control, the framework effectively transforms the complex time-varying plant into a stabilized linear system. Comparative simulations demonstrate that the proposed method achieves superior tracking fidelity compared to conventional fixed-gain baselines. Crucially, it effectively eliminates phase lag during rapid stiffness transitions and demonstrates enhanced energy efficiency, validating its capability to maintain high-precision control while preserving mechanical compliance.

GS17-2 Tracking control of autonomous underwater vehicle equipped with neural network controller optimized using genetic algorithm under rip-current environment

Tenyu Minei[1], Kunihiko Nakazono[2], Eiho Uezato[2] and Naoki Oshiro[2]
([1]Graduate school of University of the Ryukyus, Japan)
([2]University of the Ryukyus, Japan)

In this study, we propose a neural network controller optimized using a genetic algorithm (GA) for the tracking control of an autonomous underwater vehicle operating in a rip-current environment. The connection weights of the neural network are optimized using the GA. The training environment simulates a scenario in which a uniform current of constant magnitude is generated within a specific zone. The performance of the proposed controller is verified through simulation experiments via five test cases under different current conditions.

GS17-3 Implementation of an End-to-End Vision-Based Autonomous Driving System on an AI Formula Unmanned Vehicle

Wei Zhao[1], Hongkang Yu[1], Zhewen Zheng[1], Matthias Rätsch[2], Michael Danner[3] and Wenjing Cao[1]
([1]Graduate School of Science and Technology, Sophia University, Japan)
([2]ViSiR, Reutlingen University, Germany)
([3]Bochum University of Applied Sciences, Germany)

This work presents a vision-based end-to-end imitation learning approach for autonomous driving of a small-scale mobile robot developed in the Honda AI Formula project. The aim of this study is to investigate the practicality of end-to-end visual control on a real robotic platform operating under limited computational resources. A convolutional neural network based on ResNet-18 is employed to directly map front-facing camera images to continuous control commands, enabling autonomous driving without explicit perception, trajectory planning, or rule-based modules. Training data are obtained from human driving demonstrations collected on a closed test track. The trained model is deployed on an embedded computing platform and evaluated through real-world driving experiments. Experimental results on a closed course show that the proposed approach achieves smooth and stable autonomous driving, closely following the lane while maintaining robust performance. Compared with a conventional modular control pipeline, the end-to-end method demonstrates improved trajectory stability and driving efficiency. These results suggest that vision-based end-to-end control is a feasible and effective solution for small-scale autonomous driving systems and offers a promising direction for simplifying system design in real-world robotic applications.

GS17-4 Autonomous Underwater Vehicle Tracking Control Using Neural Network Controllers Optimized by GA and Enhanced by Tidal Current Estimation

Kano Kubo[1], Kunihiko Nakazono[2], Naoki Oshiro[2] and Hiroshi Kinjo[2]
([1]Graduate School of Engineering and Science, University of the Ryukyus, Japan)
([2]University of the Ryukyus, Japan)

Large-scale hydrothermal mineral deposits rich in copper and zinc have been discovered in the waters surrounding Okinawa and are estimated to be among the largest in Japan. Their development is of significant economic importance, particularly for resource-poor Japan. Although remotely operated vehicles (ROVs) are widely used for seabed surveys, their tethered cables restrict mobility, increase sensitivity to environmental disturbances, and result in high operational costs. To address these limitations, this study focuses on autonomous underwater vehicles (AUVs), which operate without cables and enable flexible, fully autonomous navigation. This paper proposes a neural network–based tracking control method for AUVs that integrates two controllers: an initial learning controller optimized offline using a genetic algorithm and an additional learning controller that updates online using estimated tidal current information. Ocean current disturbances are estimated from deviations between the target path and the actual vehicle trajectory and are incorporated into the learning process to improve control accuracy. The proposed method is evaluated through simulation in terms of adaptability to environmental disturbances, stability of learning, accuracy of current estimation, and path-tracking performance. The results demonstrate that the proposed approach effectively enhances robust and precise path-following capability in dynamic underwater environments.

GS17-5 Introduction to Robotic Welding Algorithms Using AI-Based Vision Methods

Kangmin Noh, Yunchae Oh and Hyunhwan Jeong
(Korea University, South Korea)

This paper presents an autonomous robotic welding algorithm that integrates an AI-based vision model with a 3D sensing system. The proposed method utilizes a deep learning model trained on welding images to detect and localize the welding region. Upon detection, the region of interest is mapped to corresponding 3D point cloud data. Subsequently, a convexity-based algorithm identifies workpiece edges to analyze the geometric relationship between edges and surfaces. This analysis enables precise gap detection and optimal trajectory generation. The feasibility of the proposed framework was validated through experiments using a 6-DoF robotic manipulator equipped with a 3D laser scanning system.

GS17-6 Wavelet and Neural Network-Based Diagnosis of Bearing Faults in Induction Motors under Variable Loads

Hwi Gyo Lee, Seon Min Yoo, Wang Ke Hao and In Soo Lee
(Kyungpook National University, South Korea)

This study shows induction motor bearing fault diagnosis under variable load conditions using the continuous wavelet transform. SVM, MNN, and RF classifiers are compared in terms of accuracy and computational time across three feature sets.

January 23 (Friday), 9:00-10:30

B1 Meeting Room 4

GS24 Medical informatics & Biomedical imaging

Chair: Timothée Levi (University of Bordeaux, France)

GS24-1 Blood Pressure Estimation Based on Asymmetric Sine Wave Model Residuals Optimized for PPG in 30fps Visible Camera Environment

Yusuke Nakazawa, Kent Nagumo and Akio Nozawa
(Aoyama Gakuin University, Japan)

We propose a novel blood pressure estimation method that does not depend on frequency decomposition for pseudo-PPG derived from smartphone visible cameras. In a low sampling rate environment of 30fps, conventional frequency analysis approaches struggle to accurately extract high-order harmonics. In this study, we adopt an existing morphological feature approach (RTBP: RealTimeBloodPressure) as a baseline and implement and compare two newly proposed methods. First is a linear regression model (sinBP(M): sinBP(Model)) that directly uses sin wave fit parameters (amplitude, phase, mean) as features. Second is a three-stage estimation model (sinBP(D): sinBP(Distortion)) that uses the residual (distortion index E) from a physiologically plausible asymmetric sin wave model as a feature. sinBP(M) and sinBP(D) extract essential characteristics of the PPG waveform and realize stable blood pressure estimation that is less susceptible to noise. The evaluation of this study compares three different blood pressure estimation methods (RTBP, sinBP(M), sinBP(D)) using an rPPG environment with a 30fps visible camera, with a continuous blood pressure monitor as a reference, to verify the superiority of the novel methods.

GS24-2 Unfocused-gaze indices for hypoglycemia-related lapses

Yuto Nishioka and Takehito Kikuchi
(Oita University, Japan)

Hypoglycemia can impair driving-related attention, motivating practical fixation/non-fixation markers for low-cost hardware. Many eye-tracking indices depend on temporal gaze dynamics and are sensitive to sampling rate; therefore we focused on simple spatial properties of gaze trajectories in a sensor-centered coordinate system reflecting combined head–eye orientation. Using a low-cost eye tracker (Tobii Eye Tracker 4C), we recorded gaze from 17 adults under near fixation (0.8 m), far fixation (25 m), and intentionally unfocused conditions. The midpoint of binocular 3D eye positions was projected onto a virtual plane 0.8 m in front of the sensor, with data quality controlled by IPD range checks and median absolute deviation–based outlier removal. From the projected trajectories, we extracted spatial indices reflecting concentration/spread, revisit/drift behavior, and clustering, normalized them to each participant's 0.8 m baseline condition, and compared 25 m vs. unfocused using paired Wilcoxon signed-rank tests. The unfocused condition showed significant increases in spatial entropy, convex hull area, revisit distance, and DBSCAN cluster count relative to 25 m fixation. These results support such spatial indices as robust fixation/non-fixation markers suitable for low-cost hypoglycemia risk monitoring.

GS24-3 Cross-Subject Validation of Lower Limb Joint Angle Estimation Using Single-IMU Sensor

Koyo Toyoshima and Jae Hoon Lee
(Ehime University, Japan)

Accurate joint angle estimation using minimal sensor configurations remains a critical challenge for deploying gait analysis systems in clinical and home-based settings. While previous studies have demonstrated the feasibility of single-sensor approaches, the generalizability across different individuals and the practical implementation using consumer-grade devices have not been thoroughly investigated. This study addresses these limitations through two complementary experiments. First, we evaluated the generalizability of a single-sensor joint angle estimation system across 17 healthy participants using Leave-One-Subject-Out-Cross-Validation. A Deep learning model was trained to estimate bilateral hip, knee, and ankle joint angles in the sagittal plane from a single inertial sensor attached to the pelvis. The model achieved an overall mean absolute error (MAE) of 4.126±2.813° and correlation coefficient of 0.967 when compared against reference measurements from an IMU-based motion capture system. However, detailed analysis revealed considerable inter-subject variability in estimation accuracy (MAE range: 2.155°-6.237°). Second, we validated the practical applicability by evaluating the trained model using data collected from a smartphone's built-in inertial sensor, which showed preliminary evidence of comparable performance to dedicated IMU devices. These findings provide important insights for real-world deployment of single-sensor gait analysis systems.

GS24-4 Characterizing Head Sway Responses to VR Visual Stimuli with Varying Direction and Magnitude

Tomoya Kato, Muhammad Shulhan Khairy, Wen Liang Yeoh and Osamu Fukuda
(Saga University, Japan)

Falls in older adults are a critical issue, often stemming from reduced balance ability and a heightened reliance on visual information. Traditional balance tests may overlook impairments compensated by vision. This study aimed to quantitatively assess how dynamic visual stimuli affect static standing posture using an accessible, portable Head-Mounted Display (HMD) system. We developed a VR system to present oscillating visual environments in three principal directions (Forward-Backward, Left-Right, Up-Down) at two distinct frequencies (0.4 Hz, 0.8 Hz) to healthy young participants (N=5). Our key objectives were to examine (1) the influence of stimulus direction and frequency on body sway, and (2) the correspondence between HMD-derived head position data and conventional Center of Pressure (CoP) measurements. Results demonstrated a strong correlation between CoP and head position data in the AP and ML directions, confirming the HMD's utility as an alternative body sway measurement tool. Visually induced sway was highest with FB oscillation at 0.4 Hz (in AP sway) and LR oscillation (in ML sway). UD oscillation had the least effect. This research enhances the understanding of visual control in posture and provides a foundation for developing a novel, sensitive, and portable balance assessment technique using HMD technology.

GS24-5 Multimodal EEG–ECG Framework for Dementia Recognition through Analysis of Local Neural Activity and Brain–Heart Functional Connectivity

Kusum Tara[1], Ruimin Wang[2], Yoshitaka Matsuda[3], Satoru Goto[2], Takako Mitsudo[4,5], Takao Yamasaki[4,6] and Takenao Sugi[2]
([1]Department of Biological and Material Engineering, Graduate School of Science and Engineering, Saga University, Japan)
([2]Department of Electrical and Electronic Engineering, Faculty of Science and Engineering, Saga University, Japan)
([3]Institute of Ocean Energy, Saga University, Japan)
([4]Minkodo-Minohara Hospital, Fukuoka, Japan)
([5]NHO Hizen Psychiatric Medical Center, Saga, Japan)
([6]International University of Health and Welfare, Fukuoka, Japan)

Dementia progressively disrupts psychological and physiological functioning, with alterations in both brain and heart activity reflecting its severity. This study presents a multimodal EEG–ECG framework designed to differentiate prodromal dementia (PD) from severe dementia (SD) by analyzing EEG local neural activation features, ECG autonomic regulation features, and EEG–ECG functional connectivity features. Simultaneous EEG and ECG were acquired from 10 mild cognitive impairment subjects as PD and 10 Alzheimer's disease subjects as SD. Fisher's ratio selected the sympathetic nervous system to parasympathetic nervous system (SNS to PNS) ratio and magnitude-squared coherence (MSC) as the most discriminative features during eyes-closed (EC), while hemispheric alpha asymmetry index (HI) and phase synchrony (PS) were most informative during eyes-open (EO). ANOVA further confirmed significant PD and SD differences ($p < 0.05$) across EC and EO tasks with emphasizing autonomic imbalance, MSC and PS connectivity deterioration. Coherence analysis showed consistently weakened anterior brain-to-autonomic nervous system (ANS) and posterior brain-to-ANS coupling across 0.5–30 Hz in SD, reflecting progressive breakdown of neural–autonomic integration. This proposed multimodal EEG–ECG framework effectively captures dementia-related brain–heart connectivity alterations, offering a promising strategy for assessing dementia severity and supporting monitoring the progress of dementia treatments.

GS24-6 A Multimodal Informatics-Based Evaluation of Physiological Responses to Contact Needle Technique (CNT) and Kampo diagnosis

Yihan Yuan, Hongyang Li, Yukari Uryu, Yukina Kotaka, Yuka Nagatomo, Aya Inoue, Keiko Hirose and Keiko Ogawa-Ochiai
(Hiroshima University, Japan)

Contact needle technique (CNT) is a non-invasive acupuncture method used in Kampo medicine, yet its physiological effects remain insufficiently clarified. This study aimed to objectively evaluate the immediate effects of CNT on fatigue, facial skin characteristics, and peripheral microcirculation, and to examine potential sex-related differences. Sixteen adults with fatigue received a single CNT session. Subjective symptoms were assessed using the Visual Analogue Scale, and objective measures—including facial texture and pigmentation parameters, microcirculatory indices, and heart rate variability—were obtained before and after treatment. CNT significantly reduced fatigue and improved several other subjective symptoms. Microcirculatory responses demonstrated increased superficial and deep blood-flow velocities, predominantly in women, who also showed decreases in superficial oxygen saturation and hemoglobin concentration. In contrast, men exhibited increased grayscale intensity

and a tendency toward higher facial hemoglobin values, along with sex-related differences in sulcus cutis thickness. No significant changes were observed in melanin indices or autonomic function. These findings suggest that CNT may modulate microcirculation and facial physiological features, with distinct response patterns between sexes.

January 23 (Friday), 10:45-12:00

3F Meeting Room 33

OS7 AROB: Biomimetic Machines and Robots I

Organizers: Keigo Watanabe (IPU Tokyo, Japan / Okayama University, Japan)
Kiyotaka Izumi (Saga University, Japan)
Fusaomi Nagata (Sanyo-Onoda City University, Japan)

OS7-1 Automation of Capping Task for Resin-Molded Articles by an HCLS Data-based Industrial Robot

Shingo Sakata[1], Fusaomi Nagata[1], Hayato Shimogama[1], Yota Imamura[1], Hirohisa Kato[1], Takeshi Ikeda[1], Akimasa Otsuka[1], Kiyoshi Fujisawa[2], Keigo Watanabe[3] and Maiki K. Habib[4]
([1]Sanyo-Onoda City University, Japan)
([2]Mitsubishi Pencil Co. Ltd., Japan)
([3]Okayama University, Japan)
([4]The American University in Cairo, Japan)

We investigated the automation of the capping task as another additional function for the previously developed sorting robot for defective fibrous workpieces. The capping task here refers to the task of properly attaching the resin fibrous workpiece body to a resin cap with a click sound. In the past, when a robot attempted to attach a product held by its gripper to a cap, not only it was necessary to perform delicate positioning control in the vertical direction and calibrate the robot coordinate system and workpiece coordinate system, but also the pressing load could also cause the jig to deflect, making it impossible to achieve reliable automation with just a fixed pressing operation. We therefore focused on the click sound that occurs when the product is successfully fitted into the cap, and developed a system that can perform the capping task stably and continuously by stopping the robot's pushing action when it detects this sound.

OS7-2 Training Images Searching Required for Generalization Performance of CNN Using Interactive Genetic Algorithms

Guangsheng Shao[1], Fusaomi Nagata[1], Shiori Nakashima[1], Keigo Watanabe[2] and Maki K. Habib[3]
([1]Sanyo-Onoda City University, China)
([2]Okayama University, Japan)
([3]The American University in Cairo, Egypt)

Genetic Algorithms (GAs) simulate the process of natural evolution to find optimal or near-optimal solutions to complex problems. By iteratively applying selection, crossover, and mutation operations, GAs evolve a population of candidate solutions toward better performance. GAs are widely used in optimization, scheduling, and machine learning tasks where it is difficult for traditional gradient-based methods to be applied or to solve them. This paper aims to construct a system that employs interactive GAs to attempt the automatic adjustment of such parameters as max epochs, mini-batch size, initial learning rate, and usage rate of training images, thereby proposing guiding conditions to assist designers in building desirable CNN models. By interacting with the user through our developed MATLAB application that provides the functionality of interactive GA, it becomes possible to automatically build the user's desired CNN model, e.g., for defect detection of industrial products.

OS7-3 Abnormality Detection Using Fully Convolutional Data Description Model in Metal Milling Process by CNC Machine Tools

Tomoaki Morimoto[1], Fusaomi Nagata[1], Haruto Tsubone[1], Hiroyuki Takada[1], Keigo Watanabe[2] and Maki K. Habib[3]
([1]Sanyo-Onoda City University, Japan)
([2]Okayama University, Japan)
([3]The American University in Cairo, Japan)

The authors have been developing a design, training and building application with a user-friendly operation interface for CNN (Convolutional Neural Network), CAE (Convolutional Autoencoder), SVM (Support Vector Machine), YOLO (You Look Only Once), FCN (Fully Convolutional Network) and so on, which can be used for the defect detection of various kinds of industrial products even without deep skills and knowledges concerning information technology. In those models, images are basically used for training data. Intelligent abnormally diagnosis system for computer numerical control (CNC) machine tools has been considered, i.e., what structures of neural networks should be applied. Mechanical sound and vibration generated from a machine tool itself or machining sound and vibration generated from a router bit, i.e., end mill cutter is recorded and used for training data. Extracted sound block data (SB data) from WAVE files are employed for training NN models. It has been already confirmed from preliminary experiments that a 1D CNN and an autoencoder are effective for a classification task and an identification one, respectively. In this paper, a SB data-based fully convolutional data description (FCDD) model is proposed for abnormally detection of removal machining by CNC machine tools and its concurrent visualization, in which time series data such as SB data can be directly applied to training and testing. The effectiveness of the proposed method is shown through actual experiments.

OS7-4 LSTM-based Robust Lift Detection for Autonomous Mobile Robots

Takeru Minami[1] and Shoichi Maeyama[2]
([1]Graduate School of Science for Creative Emergence, Kagawa University, Japan)
([2]Faculty of Engineering and Design, Kagawa University, Japan)

Accurate self-localization is critical for mobile robots, but it fails when a small robot is lifted, causing odometry breakdown. Prior studies using Boosting Algorithms lacked generalizability to irregular lifting motions due to limitations in capturing complex temporal dependencies in time-series data. To solve this, we propose an LSTM-based lift detection system that requires no additional hardware, maximizing the use of existing internal sensors (IMU, motor encoders). Leveraging LSTM's superior long-term pattern recognition capabilities, our deep learning approach accurately and robustly identifies diverse, irregular lifting motions. Experimental verification using an actual robot demonstrated that the proposed system exhibits superior generalizability and high real-time inference performance compared to the prior work. This improved robustness against irregular external interference makes the system highly reliable. Furthermore, we confirmed the effectiveness of a feedback control system utilizing the detection result, which immediately stops the wheels and outputs a warning, confirming an improvement in operational safety. Our ultimate goal is to enable uninterrupted self-localization even under complex disturbances.

OS7-5 An Affine Dynamic Model of an OTEC Plant Without Using Steady-State Outlet Temperatures of Heat Exchangers

Keigo Watanabe[1], Shilin Yi[2], Kiyotaka Izumi[3] and Yasuyuki Ikegami[4]
([1]Graduate School of Env., Life, and Nat. Sci. and Tech., Okayama University / International Pacific University (IPU) Tokyo, Japan)
([2]School of Mechanical Engineering, University of South China, China)
([3]Department of Mechanical Engineering, Saga University, Japan)
([4]Institute of Ocean Energy, Saga university, Japan)

Conventional models for Ocean Thermal Energy Conversion (OTEC) plants require solving nonlinear equations iteratively to obtain steady-state temperatures at the evaporator and condenser outlets, which imposes significant computational burdens in dynamic analysis and real-time control design. To address this issue, this paper proposes a novel affine dynamic model that eliminates the explicit calculation of steady-state outlet temperatures and directly represents the dynamic characteristics of heat exchangers using transient heat transfer rates and outlet temperatures. Furthermore, the turbine and pump outputs are linearized with respect to the warm and cold seawater outlet temperatures, enabling explicit expression of the plant output as a linear combination of state variables. This approach simplifies the model structure while maintaining a form suitable for control design, facilitating sensitivity analysis and robust control system design with realistic computational costs. The proposed model contributes to the dynamic characteristic analysis of OTEC plants and the development of advanced control strategies.

January 23 (Friday), 10:45-11:45

B1 Meeting Room 1

OS6 AROB: Bio-inspired Theories and Applications (2)

Organizers: Kunihito Yamamori (University of Miyazaki, Japan)
Masaru Fukushi (Yamaguchi University, Japan)

OS6-1 Investigation of an Authentication Method Using Vibration-Based Secret Cues and a Virtual Dial Interface

Takumi Kido[1], Shotaro Usuzaki[1], Kentaro Aburada[1], Hisaaki Yamaba[1], Nobuya Takahashi[1], Mirang Park[2] and Naonobu Okazaki[1]
([1]University of Miyazaki, Japan)
([1]Kanagawa Institute of Technology, Japan)

Personal authentication on mobile devices, such as PINs, is vulnerable to shoulder-surfing attacks where attackers observe screen taps. While previous studies introduced vibration-based cues to hide inputs, they often required burdensome full-dial rotations, leading to poor usability. This paper proposes an improved authentication method resistant to shoulder-surfing attacks using device vibration and a virtual dial interface. In this scheme, the initial vibration position is randomized and known only to the user. Users input passcodes by rotating the dial from this secret vibration position by the number of steps corresponding to the passcode digit. Consequently, an attacker observing the screen can only see the total rotation steps, preventing them from deducing the specific code. To evaluate the method, we conducted experiments with 16 participants using three different visual dial patterns. The results showed an average authentication time of 16.7 seconds and an authentication success rate of 79.2%. Crucially, the shoulder-surfing success rate was 0%, demonstrating robust security. The System Usability Scale (SUS) score was 69.5, indicating that the method ensures sufficient usability while maintaining high security.

OS6-2 Malware Classification Using Transformer-Based Modeling of API Call Sequences

Ryoga Sakai[1], Shotaro Usuzaki[1], Kentaro Aburada[1], Hisaaki Yamaba[1], Nobuya Takahashi[1], Mirang Park[2] and Naonobu Okazaki[1]
([1]University of Miyazaki, Japan)
([2]Kanagawa Institute of Technology, Japan)

Cyber threats posed by malware are becoming increasingly sophisticated, making accurate and efficient malware analysis essential. Traditionally, Application Programming Interface (API) call sequences have been treated as time-series data, and classification methods based on Recurrent Neural Networks(RNNs) and Long Short-Term Memory(LSTM) networks have been widely adopted. However, these approaches struggle to capture long-range dependencies due to the vanishing gradient problem. In contrast, Transformer-based models can efficiently model long-range dependencies across entire sequences. Although API call sequences are typically long, they also show behavioral diversity even within the same category. As a result, the choice of model architecture and preprocessing strategies can significantly influence classification performance, yet this impact has not been systematically examined. This study compares four models—RNN, LSTM, Transformer, and Bidirectional Encoder Representations from Transformer (BERT)—under unified experimental conditions for malware classification using API call sequences. Experimental results show that the Transformer model achieves stable performance across many categories, whereas RNNs and LSTMs show performance variations depending on sequence length and category characteristics. Furthermore, performance differences across categories are observed for BERT due to variations in tokenization methods, highlighting the importance of designing preprocessing strategies tailored to API call sequences.

OS6-3 A Browser-based Parallel VC Client using WebAssembly

Keiichi Inohara, Sun Zhiding, Yota Kurokawa and Masaru Fukushi
(Yamaguchi University, Japan)

Volunteer computing (VC) is one of the distributed computing paradigms, which constructs a distributed computing environment on the Internet. Nodes in VC operate on the open Internet and thus neither recognize other workers nor accept connections from them. In other words, workers are typically not configured to communicate with others. Hence, application of the current VC is limited to ``bag-of-tasks" computations, and parallel computations that require data exchange between workers are not supported. In our previous study, toward the realization of a VC that supports parallel computations, we have developed a parallel VC system based on the concept of server-assisted communication. Through a verification experiment, we have confirmed the correct operation of the server-assisted communication mechanism and the execution of several parallel benchmark programs, demonstrating the feasibility of parallel VC. However, since each worker was virtually implemented as a Linux process for the purpose of verifying the operation of parallel VC, the worker environment has not yet been implemented. In this paper, we propose a browser-based VC client using WebAssembly for the parallel VC system. We evaluated execution time of the proposed client using Integer Sort from the NAS Parallel Benchmarks. The evaluation results showed that the proposed method incurred an overhead of approximately five seconds, which is acceptable for practical VC use.

OS6-4 Cracker Identification from the Embedded MAC address into TCP Header

Soma Ohashi[1] and Kunihito Yamamori[2]
([1]Graduate School of Engineering, University of Miyazaki, Japan)
([2]Faculty of Engineering, University of Miyazaki, Japan)

With the continued expansion of Internet usage, cyberattacks have become increasingly frequent and sophisticated. In particular, IP spoofing and botnet-based distributed denial-of-service (DDoS) attacks pose serious challenges to servers and network administrators because attackers can conceal their identities and generate traffic at massive scale. Conventional defense mechanisms mainly rely on IP addresses to identify and block malicious hosts; however, such approaches are often ineffective against spoofed or frequently changing IP addresses. This study proposes a novel server protection method that leverages the extensibility of the TCP/IP protocol by embedding Media Access Control (MAC) addresses into the TCP header. By incorporating MAC address information into connection request packets, the server can identify attacking terminals more accurately and maintain a denylist independent of IP addresses. Once a malicious terminal is detected, its MAC address is registered and subsequent connection attempts from the same source can be rejected at an early stage. This approach aims to suppress repeated attacks from identical hosts, reduce unnecessary processing overhead, and enhance resilience against large-scale DDoS attacks.

January 23 (Friday), 10:45-12:00

B1 Meeting Room 2

OS23 AROB: Toward Secure, Resilient, and Intelligent Autonomous Decentralized Control Systems

Organizer: Kenji Sawada (The University of Osaka, Japan)

OS23-1 Adaptive Kalman Filter for Sensor Importance Evaluation and Safe Continuation of Operation

Kaito Higashi, Yuichiro Sueoka and Kenji Sawada
(The University of Osaka, Japan)

Reliable localization in autonomous robots requires robust fusion of heterogeneous sensors, yet practical operation often suffers from packet loss and intermittent observations. Such degradation reduces the information available to the estimator and can lead to unsafe behavior if the system continues operating without recognizing loss of sensing quality. This paper presents an Adaptive Kalman Filter (AKF) that adjusts the diagonal elements of the process noise covariance Q according to each sensor’s packet loss rate. The trace of the estimation error covariance, trace(P), is used as a unified uncertainty indicator, enabling quantitative assessment of sensing degradation and a threshold-based decision on safe continuation. The method was implemented by extending the EKF in ROS 2 robot_localization and evaluated on a TurtleBot4 under controlled IMU and odometry dropout. Odometry loss caused a sharp increase in trace(P) and path-tracking failure, whereas IMU loss produced only mild effects. A threshold on trace(P) allowed the robot to stop safely before trajectory collapse. Sensitivity analysis of ∂trace(P)/ ∂Q_s further revealed the estimator’s dependence on individual sensors. These results demonstrate that the proposed framework

provides both a reproducible safety criterion and a practical tool for analyzing sensor importance under degraded sensing.

OS23-2 Design and Implementation of Visualization Functions in a Cybersecurity Exercise Program for Control Systems

Masatoshi Enomoto[1], Shu Hosokawa[2] and Kenji Sawada[3]
([1]Yokohama College of commerce, Japan)
([2]Seisho PWB Design Office, Japan)
([3]The University of Osaka, Japan)

This study proposes and evaluates a visualization system designed for industrial control system (ICS) cybersecurity exercises, specifically targeting operational technology (OT) operators with limited IT knowledge. While OT operators can easily observe physical process anomalies during a cyberattack, identifying the underlying network causes remains a significant challenge. To bridge this gap, we implemented a visualization tool based on three core requirements: (1) information volume reduction, (2) identification of critical protocol layers, and (3) translation of machine-to-machine (M2M) communications into human-readable formats through topological mapping. The system was evaluated using a low-cost testbed across three attack scenarios: Denial of Service (DoS), Replay, and Man-in-the-Middle (MitM). Experimental results demonstrated that while DoS attacks are easily identified by traffic volume, MitM attacks are highly stealthy, showing a minimal packet reduction rate of only 1.96% after filtering. However, by providing topological visualization, the proposed system allows beginners to intuitively recognize unauthorized structural changes in communication paths. These findings confirm that the system effectively lowers the cognitive load for beginners and supports the ”Check” phase of the PDCA cycle in incident response training.

OS23-3 Experimental Evaluation of Human-in-the-loop Platooning Control

Shunsuke Horiuchi, Kaito Higashi and Kenji Sawada
(The University of Osaka, Japan)

Adaptive Cruise Control (ACC) and Cooperative Adaptive Cruise Control (CACC) are critical autonomous driving systems that smooth traffic flow by adjusting following distances. CACC uses vehicle-to-vehicle (V2V) communication to improve safety and coordination, but its effectiveness is limited in mixed traffic environments with both autonomous and human-driven vehicles, potentially reducing control performance and traffic capacity. This study experimentally investigates platoon control incorporating human-driven vehicles into the control loop, evaluating the impact of communication and computational delays on control performance. A platoon consisting of three vehicles was constructed using Raspberry Pi Mouse V3 robots, with the first and third vehicles operating autonomously and the second vehicle manually. The control architecture combined PID control with inter-vehicle distance control, with the third vehicle estimating distance using a camera. Experimental results demonstrated that CACC-based control effectively avoided collisions even during sudden acceleration by the human-driven vehicle. However, communication and computational delays caused response delays and amplified speed errors, revealing that complete platoon stability was not achieved. Future research will focus on thoroughly elucidating the effects of these delays.

OS23-4 Online Optimization of Pickup and Delivery Problems Considering Accident Recovery Forecast

Soma Nakagawa, Koichi Kobayashi and Yuh Yamashita
(Hokkaido University, Japan)

In this paper, online optimization of the pickup and delivery problem is studied. In particular, we focus on incorporating an accident recovery forecast. First, the target area is modeled by a set of graphs. Next, constraints and cost functions are introduced. Then, the pickup and delivery problem is rewritten as a mixed integer linear programming problem. Finally, we provide numerical examples illustrating the computational results under accident scenarios with the incorporation of an accident recovery forecast.

OS23-5 A Data-Driven Design Algorithm of Prediction Governors

Yuki Minami
(University of Hyogo, Japan)

This paper focuses on the design problem of prediction governors, which are add-on modules designed to improve the reliability of control systems using prediction signals generated by AI models. In our previous studies, we developed a model-based optimal design method for prediction governors. However, the model-based approach has practical limitations, such as the need for accurate system models. To overcome these challenges, we propose a data-driven design algorithm that constructs prediction governors directly from input-output data collected from the system, without requiring explicit model knowledge. In this paper, we first formulated the design problem of prediction governors, and provided a brief review of the optimal model-based design method. Next, the data-driven design algorithm was developed. This method eliminates the need for an explicit system model and guarantees a low-order structure, thereby reducing the computational cost. Finally, the effectiveness of the proposed method was demonstrated through a numerical simulation.

January 23 (Friday), 10:45-12:00

B1 Meeting Room 3

GS16 Identification and Estimation

Chair: Yuichiro Taira (National Defense Academy, Japan)

GS16-1 Estimation of Lower-Limb Joint Angles during Various Gait Patterns Using a Single Waist-Mounted Inertial Sensor and Deep Learning

Masato Yabe, Koyo Toyoshima and Jae Hoon Lee
(Ehime University, Japan)

Currently, gait abnormalities in patients are identified through medical interviews, visual observation of gait, and numerical evaluation using simple tests. However, these methods have problems such as lack of reproducibility, lack of objectivity, and the risk of overlooking abnormalities. Furthermore, optical and inertial motion capture systems are difficult to use in clinical settings due to the large amount of equipment required, the burden on patients due to space limitations, and high cost. Therefore, previous research has used deep learning to estimate joint angles from acceleration, angular velocity, and data from an IMU (inertial sensor) attached to the waist. While the results were highly accurate and favorable, the data only included healthy subjects, so the applicability to diverse gaits, such as those of elderly and patients, was unknown. Therefore, this study aimed to estimate lower limb joint angles during diverse gaits. As a result, although there were some areas where the joint angle estimation graph lacked accuracy, the overall waveform shape was reproduced for all gaits. Therefore, it was demonstrated that by using a variety of walking data, joint angle estimation is possible even for complex gaits.

GS16-2 Study on Joint Angle Estimation Accuracy During Walking Motion Based on Sensor Placement Differences

Tatsuya Ebisu and Jae Hoon Lee
(Ehime University, Japan)

Physicians typically assess gait abnormalities through visual observation, but this evaluation relies heavily on their experience and subjective judgment, making the lack of objective metrics a significant limitation. While optical and inertial sensor-based motion capture systems are established methods for measuring gait, both approaches are costly and require substantial setup time. To address this, a previous study used deep learning to estimate joint angles from acceleration and angular velocity data obtained from an IMU (inertial measurement unit) worn at the waist. Building on this work, the present study investigated whether placing sensors on body parts other than the waist could improve gait estimation efficiency. Participants wore both an IMU-based motion capture system (Xsens) and an inertial sensor (E2box). Sensors were placed at three locations: the waist and both feet. Participants walked normally along a 6-meter path for five minutes. During walking, joint angles were recorded using Xsens, while acceleration and angular velocity were recorded via the E2box. Data from six participants were used to train a convolutional neural network (CNN). The CNN took acceleration and angular velocity from three input cases: waist only (Case 1), feet only (Case 2), and waist plus feet (Case 3). The network estimated angles of six joints: both hips, knees, and ankles.

GS16-3 Phase-Difference Based Estimation of Internal Defect Directions in Infrastructure Structures

Kenshin Yamakawa, Geunho Lee, Hayato Kawazu and Keito Katsuki
(University of Miyazaki, Japan)

In Japan, many concrete infrastructures built during the period of rapid economic growth are now deteriorating, creating an urgent need for efficient inspection methods. Although technologies such as ultrasonic testing, impact acoustic testing, and ground-penetrating radar have been introduced, they require expensive equipment and specialized expertise. Therefore, percussion-based inspection remains the most widely used method in practice. However, conventional impact inspections mainly provide one-dimensional information, such as the presence or depth of defects, and few studies have focused on improving overall efficiency. This study proposes a new method for estimating the direction of internal defects in concrete structures to enhance inspection efficiency. The approach identifies and extracts reflected waves generated by an impact and estimates their angle using two microphones. Because reflected waves primarily appear in high-frequency bands, frequency-domain analysis is used to isolate them. A geometric model is then developed to calculate defect direction based on the difference in arrival distance between the microphones. Experiments using specimens with artificial defects verify the validity of the method, demonstrating successful extraction of reflected waves and estimation of defect direction. This approach is expected to improve inspection efficiency and, in the future, support unmanned tunnel inspection through integration with robotic systems.

GS16-4 Development of path generation algorithm for forest survey multicopters considering measurement coverage

Rui Ozawa and Masami Iwase
(Tokyo Denki University, Japan)

This study aims to contribute to carbon neutrality and promote forestry by developing a forest survey system using a LiDAR-equipped multicopter for high-precision 3D mapping. An algorithm is proposed to automatically determine optimal measurement positions in order to improve the measurement coverage rate, defined as the proportion of trees measured within a survey area. Unlike conventional forest surveys, where measurement positions are selected based on the experience and intuition of forestry workers, the proposed method systematically determines suitable positions using LiDAR point cloud data. Initial LiDAR measurements are conducted from predefined positions along forest roads to roughly estimate tree locations and radii in an unknown forest, and a two-dimensional forest map is generated. Based on this map, multiple candidate measurement positions are evaluated, and the position that captures the largest number of trees is selected. The forest map is then updated, and this process is repeated iteratively. Furthermore, a deep neural network–based inference model is introduced to accelerate measurement position selection and reduce computational cost. Simulation results demonstrate that the proposed method improves measurement coverage and processing speed, enabling practical application to autonomous forest survey multicopters.

GS16-5 Wi-Fi CSI-Based Abnormality Detection in Private Toilets Using Posture Estimation and Stillness Determination

Kenta Yazu[1] and Atsuo Ozaki[2]
([1]Graduate School of Information Science and Technology, Osaka Institute of Technology, Japan)
([2]Department of Information and Computer Science, Osaka Institute of Technology, Japan)

This study aims to detect abnormal situations of users in private toilet stalls using Channel State Information (CSI) obtained from existing Wi-Fi infrastructure. We propose a method that estimates seven posture types by applying principal component analysis (PCA) and a random forest classifier to subcarrier amplitudes, and then determines stillness or movement from short-term fluctuations of the amplitudes to perform temporal smoothing of posture probabilities and obtain the final posture. Using data from three subjects, we conducted a leave-one-subject-out evaluation to assess generalization performance. As a result, the overall posture estimation accuracy improved from 0.528 (primary estimation) to 0.576 (final estimation), and the accuracy of normal/abnormal (binary) classification improved from 0.743 to 0.785. On the other hand, the accuracy remained low for a small-bodied subject, indicating that robustness against body-size differences and improvement of the primary posture estimation remain as future challenges.

January 23 (Friday), 10:45-11:45

B1 Meeting Room 4

GS30 Neurocomputing technologies and its application for hardware

Chair: Atsushi Masumori (The University of Tokyo, Japan)

GS30-1 Real time embedded spiking neural network for unsupervised biological pattern recognition

Landry Bailly[1], Jeremy Cheslet[1,2], Yoshiho Ikeuchi[2] and Timothée Levi[1]
([1]IMS, University of Bordeaux, France)
([2]IIS, the University of Tokyo, Japan)

Energy-efficient alternatives are needed for large-scale machine learning. Neuromorphic approaches, combining Spiking Neural Networks (SNNs) and bio-hybrid systems, show promise for higher energy efficiency. In this work, we present an FPGA-based implementation of an unsupervised Spiking Neural Network for biological pattern recognition. The model consists of a single layer of Leaky Integrate-and-Fire (LIF) neurons fully connected to the input with synapses following the Spike-Timing-Dependent Plasticity (STDP) rule, capable of recognizing repetitive hidden patterns. The implementation allows real-time computation (with inputs sampled at 20 kHz) and can be scaled to higher dimensions. The implementation has been validated by performing several experiments and comparing results with FPGA implementations and Python simulations. We also compared precision and recall using different number representations. This work is a part of a global project aiming to create a bio-hybrid closed loop and would be useful to processes in real time biological signals recorded with a Micro Electrode Array (MEA) to generate a feedback control.

GS30-2 Device Implementation of the Peripheral Auditory System Model

Satoshi Maebeya, Nobuo Iwasaki and Kazuya Okamoto
(National Insutitute of Technology, Wakayama College, Japan)

This study addresses the need for biologically realistic auditory input in central nervous system (CNS) models, particularly for simulations of language learning that depend on spike-based neural activity. Because CNS information processing operates through discrete spikes, auditory inputs must accurately reflect peripheral auditory transduction. Many existing models lack such inputs, limiting their ability to reproduce essential aspects of speech and language acquisition. To solve this problem, we developed a software-based input device that functions as a peripheral auditory front end for CNS simulations. The device is based on the equivalent-circuit auditory model of Giguère and Woodland, further refined using theoretical insights from Maki. By discretizing this model and implementing an updated inner-hair cell model, we achieved real-time generation of spike trains corresponding to sound pressure signals. The system converts acoustic inputs into spike sequences that mimic auditory nerve firing, enabling direct integration with CNS models. Experimental evaluation shows that the model used generates biologically valid spike patterns across various sound pressure levels and successfully captures the critical acoustic-to-neural transduction. This work enhances the biological plausibility of auditory stimulation in CNS simulations and provides a valuable tool for studying neural mechanisms of language learning and developing biologically inspired learning algorithms.

GS30-3 Field Test of Wild Boar Capture System Using YOLO Algorithm and Side-View Image

Ryunosuke Niimoto, Francis Besala and Jae Hoon Lee
(Graduate School of Science and Engineering, Japan)

In Japan, crop damage caused by wild boars has become increasingly serious, while conventional trapping methods are often large, expensive, and inefficient. To address this issue, this study proposes a compact and cost-effective automatic trapping system based on deep learning–based object detection. First, a data collection device was developed using an inexpensive microcontroller capable of running artificial intelligence and small sensors, and field data such as images of wild boars entering traps were collected. Then, a lightweight model was constructed using the You Only Look Once (YOLO) algorithm trained on the collected dataset and implemented on the microcontroller, enabling the development of a recognition device comparable in size to a conventional trail camera. Furthermore, by integrating this recognition device with the trapping device and box trap developed in our previous research, we realized a compact and low-cost automatic trapping system that activates only when wild boars enter the trap. Field experiments confirmed that the proposed system can accurately identify wild boars inside the trap and automatically

activate the capturing mechanism. This study demonstrated the potential of providing an affordable and practical solution for mitigating crop damage.

GS30-4 A Simplified Retina-Inspired Analog-Digital Motion Detection Circuit

Thanaphat Imkrajang, Hikari Fudeyasu and Kimihiro Nishio
(National Institute of Technology, Tsuyama College, Japan)

This paper presents simplified analog-digital motion detection circuits inspired by the vertebrate retina. The proposed unit circuits are designed with a focus on structural simplicity, enabling flexible arrangement in one-dimensional arrays. This configuration allows detection of both motion direction and velocity. The circuits were evaluated using the Simulation Program with Integrated Circuit Emphasis (SPICE), and simulation results demonstrated their ability to accurately detect object velocity. A Large Scale Integration (LSI) chip was fabricated using a 0.6 μm Complementary Metal Oxide Semiconductor (CMOS) process to verify the circuit operation. Measurement results confirmed that the fabricated circuit successfully detected both edge positions and motion signals. These findings indicate that compact, low power image processing sensors can be realized by applying the proposed retina-inspired circuit design.

January 23 (Friday), 13:00-14:00

3F Meeting Room 32

OS4 AROB: Applications of Deep Learning and Robotics

Organizers: Nobuya Takahashi (University of Miyazaki, Japan)
Masahiro Yokomichi (University of Miyazaki, Japan)

OS4-1 A Closed-Loop Neurofeedback System Based on Four-Channel EEG and an Abstract Avatar for Affective State Regulation

Renta Kokubo, Nobuya Takahashi and Masahiro Yokomichi
(University of Miyazaki, Japan)

Recent EEG-based neurofeedback studies have mainly focused on visualizing neural states or presenting target values, leaving state regulation largely dependent on users' self-control. This study proposes a closed-loop neurofeedback system that combines lightweight affective state estimation from low-channel EEG with an abstract avatar in a virtual space. Using four EEG channels, affective indices related to valence, arousal, and attention are estimated in real time based on spectral features in the alpha and beta bands. The estimated states are continuously reflected as dynamic changes in the avatar, providing implicit, non-instructional feedback. Evaluation results suggest that the assistive feedback condition facilitates earlier convergence toward the target affective state. These findings demonstrate the feasibility of closed-loop affective neurofeedback using minimal EEG configurations and abstract avatar-based feedback.

OS4-2 Modified Linear Attentions for Image Classification with Mobile ViT

Masahiro Yokomichi and Nobuya Takahashi
(University of Miyazaki, Japan)

Recently, many Vision Transformer (ViT) models adopts dot-product (Softmax) attention module, because of its ability to capture global visual information. However, it needs high computational cost in the time and memory consumption, thus more light-weight attention module has been investigated. The linear attention is one of them and the computational cost can be reduced by means of kernel method. Nonetheless, the worse performance limits its application domain. In this paper, two types of modifications of linear attention are proposed. The first is to use the kernel (activation) function that takes positive and negative values. This aims to approximate the similarity property of dot-product attention. The second is to combine the spatial and channel attention in order to improve the modeling performance. The effectiveness of the proposed methods are evaluated by image classification experiment with small image datasets.

OS4-3 Fixed-Point Landing of a Bird-Inspired Flapping-Wing Robot Using Model Predictive Control

Yasunobu Hitaka
(Kitakyushu National Collage of Technology, Japan)

This paper investigates the fixed-point landing (perching) capabilities of a bird-inspired flapping-wing robot using Model Predictive Control (MPC). While ornithopters offer exceptional agility in cluttered environments, achieving precise landing in confined spaces remains a significant challenge due to their complex, non-linear dynamics. To realize a soft landing with significant deceleration, the robot must execute a high-alpha "flare" maneuver just before touchdown, mimicking biological avian flight behavior. We developed a two-dimensional longitudinal dynamic model incorporating non-linear aerodynamic coefficients and formulated a constrained optimization problem. The MPC framework is designed to minimize terminal velocity and positioning errors while respecting strict physical constraints, such as elevator torque and mechanical actuator limits. Numerical simulations compared a conventional statically stable configuration with a proposed high-control-authority configuration. The analysis reveals that traditional designs prioritizing passive stability often fail to generate the necessary pitching moment for rapid deceleration. Conversely, the results demonstrate that extending the tail arm and reducing the static stability margin are essential for securing sufficient control authority. The proposed model successfully reduced the horizontal velocity to 0.894 m/s at the target point with a pitch angle of 86.97 degrees, providing a robust design and control strategy for autonomous perching in bio-inspired robots.

OS4-4 Multi-Label Classification of Spectrogram Features for Rhythm Game Chart Generation Using Recurrent Neural Networks

Taichi Inoue, Nobuya Takahashi, Masahiro Yokomichi, Shotaro Usuzaki and Kentaro Aburada
(University of Miyazaki, Japan)

Music games have become widespread across various regions of the world and have developed into diverse forms as a game genre supported by multiple generations. The chart is a crucial element that directly influences the gameplay experience and the player’s sense of immersion. It is structured to allow players to envision the distinctive melodies and rhythms used within the musical track. In general, charts are created by chart designers, and their content strongly reflects the individuality of the creator. For ordinary users, creating attractive charts is a significant burden. To address this challenge, various research efforts have explored automatic chart generation. We propose a chart generation framework that utilizes the spectrogram representation of music as the input feature. A spectrogram expresses both the frequency and time domain characteristics of the audio, enabling a richer and more detailed description of the musical structure compared to conventional methods. In addition, a multi-label classification approach is employed to provide more flexibility in chart labeling. Unlike the one-hot representation used in single-label models, this method enables the model to assign multiple actions as ground truth labels to a single time period, thereby achieving high flexibility and accuracy in representing complex musical patterns. The experimental results indicate that the distribution of action types in the generated charts closely approximates that of the training charts, suggesting that the model effectively learns and reproduces the statistical patterns of action types from the training data. The results of this study suggest that using rich audio features and probabilistic multi-label classification can greatly improve the performance of rhythm game chart generation models. These findings provide a foundation for a rhythm game content generation system based on musical knowledge.

January 23 (Friday), 13:00-14:00

3F Meeting Room 33

OS8 AROB: Biomimetic Machines and Robots II

Organizers: Keigo Watanabe (IPU Tokyo, Japan / Okayama University, Japan)
Kiyotaka Izumi (Saga University, Japan)
Fusaomi Nagata (Sanyo-Onoda City University, Japan)

OS8-1 Embedding machining errors into Topology-Optimized ideal features based on the level set method for generating skin model shapes

Yuma Hino, Fusaomi Nagata and Akimasa Otsuka
(Sanyo-Onoda City University, Japan)

Topology optimization is widely used to design lightweight and high-strength structures in fields such as automotive and aerospace engineering. However, the optimized shapes are idealized and do not consider inevitable machining errors that arise during manufacturing, which may lead to quality problems in mass production. This study proposes a method to add realistic shape deviations to topology-optimized features using the level set method combined with random surface errors. A zero isosurface is modified by applying two-dimensional normally distributed random fields, which are controlled through inverse Fourier transformation using filter width and magnification parameters. These parameters allow the scale and amplitude of deviations to be adjusted flexibly. We have confirmed that realistic variations can be introduced while maintaining optimization constraints using data generated by this method. Furthermore, actual data of surface machined by NC wire electrical discharge machining is integrated into the computational model. Measured surface data is combined with the random deviation method to generate a three-dimensional skin model that reflects real manufacturing variations more accurately. The proposed approach contributes to improving tolerance design and quality prediction for mass production.

OS8-2 Barrier Lyapunov Function-Based Autonomous Navigation of a Differential-Drive Robot Using 3D Depth Sensing

Taro Shibanoki and Keigo Watanabe
(Okayama University, Japan)

This paper proposes an obstacle avoidance behavior model for achieving autonomous navigation control of a differential-drive mobile robot designed for operation in hospital environments. The proposed method integrates 3D depth sensing based on the fusion of RGB-D and stereo cameras with object recognition. In this study, the position of objects within the robot's field of view is estimated in real time using an object detection model applied to RGB video frames. Depth information obtained from a TOF camera or similar sensor is fused with the RGB data to estimate the relative distance and orientation to the target object through sensor fusion. Based on these estimates, a nonlinear control law employing a barrier Lyapunov function (BLF) is designed to generate translational and angular velocities, enabling the robot to approach the target while maintaining a safe distance from obstacles. Simulation studies were conducted assuming multiple static and moving obstacles surround- ing the robot. The results demonstrate that the robot successfully reaches the target waypoint while avoiding obstacles using the proposed method. Although slight orientation fluctuations were observed when obstacles approached the robot closely, the robot continuously adjusts its translational and angular velocities according to the relative distance and angle, and the blending coefficient enables smooth and continuous control transitions.

OS8-3 Proposal of a 3D Pose Estimation Method Using Procrustes Analysis Based on 2D Skeletal Information

Daisuke Toki and Shogo Nonaka
(National Institute of Technology, Tsuyama College, Japan)

This paper proposes a lightweight fall detection method based on a two-dimensional skeleton of the upper body obtained from a monocular camera. Conventional methods often rely on full-body skeleton analysis or deep learning models, requiring large-scale labeled datasets and incurring high computational costs. In contrast, the proposed method achieves posture estimation based solely on 2D skeleton information, without requiring high-load learning such as depth cameras or 3D skeleton data. This is accomplished by constructing a 3D pose estimation method using a Procrustes analysis algorithm that optimizes 2D shapes. To enable handling of seated states and frame-outs, the target model defines a 3D skeleton model for the upper body only. First, this model is virtually rotated in 3D space, and its orthographic projection is optimized using Procrustes analysis. Next, the 3D pose is estimated by comparing the projected shape with the observed 2D skeleton shape using similarity evaluation, and searching for the rotation angle that minimizes the projection error. This enables estimation of body orientation without requiring estimation of the 3D coordinates of each skeleton or camera parameters. This paper establishes these algorithms as a foundation and confirms the effectiveness of 3D pose estimation through experiments.

OS8-4 Various Aspects of Interconnected Affine Models in Robot Systems --- An Approach from Heat Exchanger Models to Multi-Robot Systems ---

Keigo Watanabe[1], Shilin Yi[2], Kiyotaka Izumi[3] and Yasuyuki Ikegami[4]
([1]Graduate School of Env., Life, and Nat. Sci. and Tech., Okayama University /
International Pacific University (IPU) Tokyo, Japan)
([2]School of Mechanical Engineering, University of South China, China)
([3]Department of Mechanical Engineering, Saga University, Japan)
([4]Institute of Ocean Energy, Saga university, Japan)

This study focuses on the fact that the novel dynamic model proposed for OTEC plant heat exchangers possesses a bidirectionally coupled affine structure, where outlet temperatures and heat transfer rates are mutually interconnected, and the manipulated variables (warm/cold seawater flow rates and their reciprocals) exhibit inverse relationships. We demonstrate that this seemingly specialized structure is also commonly observed in dynamic models within robotics, including coordinated systems, multi-robot systems, and mechanisms with symmetric structures, and we systematically examine them by categorizing them into three groups. Clarifying such structural similarities not only deepens interdisciplinary model understanding but also holds significant importance in potentially establishing common approaches in future nonlinear control theory.

January 23 (Friday), 13:00-14:15

B1 Meeting Room 1

GS5 Artificial intelligence III

Chair: Reiji Suzuki (Nagoya University, Japan)

GS5-1 Applying Reinforcement Learning to Dynamic Difficulty Adjustment in Endless Games

Shunlei Xu[1], Koichi Moriyama[1], Atsuko Mutoh[1], Kosuke Shima[1], Tohgoroh Matsui[2] and
Nobuhiro Inuzuka[1]
([1]Nagoya Institute of Technology, Japan)
([2]Chubu University, Japan)

Many studies have been conducted to maintain player engagement in video games. It has been found that when game difficulty is aligned with the player's skill level, players experience enjoyment and continue playing. Therefore, dynamically adjusting game difficulty to match the player's skill level has become an important research topic. At present, there is a substantial body of research on dynamic difficulty adjustment. However, only a few studies have applied dynamic difficulty adjustment to endless games. This paper aims to investigate a dynamic difficulty adjustment method tailored to endless games. We proposed a reinforcement learning-based framework to address the dynamic difficulty adjustment problem in endless games. Through comparative experiments, the proposed method demonstrates more stable behavior and smoother adjustments than the other methods.

GS5-2 Home Garden Management System Using AI and IoT

Aka Thomas Manouan, Boburkhonov Asliddin and Hironori Hiraishi
(Ashikaga University, Japan)

In order to help farmers and particulars to monitor their garden, this study aims to develop a management system that combines AI and Internet of Things (IoT) to support sustainable home gardening, to contribute to reducing environmental impacts, and to allow farmers and ordinary people to save time. In our laboratory, various experiments were conducted using several tools to develop the system. This system is developed by using a soil moisture sensor, a small water pump, an LED light, and a light sensor, all connected to an Arduino to acquire environmental data in real time and automatically control it with an AI model by adjusting water supplied and light intensity. Image analysis using You Only Look Once (YOLO)v8 enables automatic determination of plant growth status and the presence of pests and diseases. To be more efficiency, an additional machine learning program was built to predict the precise time of the next watering.

GS5-3 Depth Estimation of Panorama Image Using Monocular Camera and Tiny Distance Sensor for Recognizing Confined Indoor Space

Kenta Kikuno, Teklay Asmelash and Jae Hoon Lee
(Graduate School of Science and Engineering, Ehime University, Japan)

In recent years, the use of drones for indoor inspections has gained significant attention. Indoor environments are characterized by confined spaces and numerous obstacles, necessitating collision avoidance technology to protect the aircraft. Furthermore, manual flight in tight spaces is hazardous, making the development of algorithms for real-time exploration and navigation essential. This research focuses on recognizing such confined spaces for a real-time exploration and navigation algorithm. We propose an algorithm that combines monocular depth estimation technology with a compact distance sensor to achieve 360° horizontal exploration and navigation. The proposed method was tested in various indoor environments, its accuracy was evaluated by comparing it with LiDAR data. The effectiveness and practicality of the proposed algorithm were confirmed through experiments.

GS5-4 Steel Surface Defect Detection Based on Improved YOLOv11 for Robotic Visual Inspection

Xinyang He and Kenji Hashimoto
(Graduate School of Information, Production and Systems, Waseda University, Japan)

In industrial strip steel production, surface quality directly affects the material's performance and reliability. Defects such as cracks, inclusions, and scratches not only weaken structural strength but may also lead to failures during downstream processing. Traditional manual inspection and fixed-position vision systems often suffer from high labor intensity, poor adaptability, and limited capability in handling complex surface textures. To meet the visual perception requirements of autonomous inspection robots in intelligent manufacturing, this paper proposes CM-YOLOv11, an improved steel surface defect detection model that achieves high accuracy, lightweight design, and real-time performance. Built upon the YOLOv11 framework, the proposed model integrates the CMUNeXt backbone module, which employs depthwise separable and pointwise convolutions to achieve efficient multi-scale feature extraction while reducing computational overhead. Furthermore, a MMSA multi-scale self-attention block is introduced in the neck to enhance global context modeling and fine-grained feature representation, thereby further improving detection performance. Experimental results on the NEU-DET dataset show that, using YOLOv11-n as the baseline, CM-YOLOv11 achieves the mAP@50 of 82.2% under a 416×416 input resolution, outperforming the original YOLOv11 (80.2%). With its favorable balance of accuracy, efficiency, and lightweight architecture, CM-YOLOv11 can be effectively deployed on industrial inspection robots and edge devices, providing a practical, efficient, and scalable solution for automated steel surface quality inspection in intelligent manufacturing.

GS5-5 A Study on Parameter Optimization for the Development of an AI Pillow that Learns Individual Sleep Patterns

Koga Sato and Shudai Ishikawa
(Department of Information Engineering, Oita National College of Technology, Japan)

Pillow height is widely recognized as a crucial ergonomic factor for maintaining good sleep quality. As humans unconsciously change posture during sleep, dynamically adjusting pillow height is essential to maintain a comfortable posture and ensure appropriate respiratory function. However, conventional static pillows fail to adapt to individual differences or tossing and turning, while existing robotic devices may negatively affect sleep quality due to physical restraint. Therefore, this research aims to develop an AI pillow that automatically adjusts to the optimal height in a non-invasive manner according to the user's state. Elucidating physical parameters, such as pressure sensor arrangement, is critical for this realization, but physical prototyping imposes significant time and cost constraints. To address these constraints, we constructed a physical simulation environment using Unity. This paper evaluates the simulator for verifying and analyzing these physical parameters. By reproducing head-pillow interactions, we aim to streamline the physical device's development and establish a foundation for control algorithms.

January 23 (Friday), 13:00-14:30

B1 Meeting Room 2

GS32 Robot vision and image processing

Chair: Abhijeet Ravankar (Kitami Institute of Technology, Hokkaido, Japan)

GS32-1 Two-Stage Point-Cloud Registration Unifying RGB-D Color Cues and Closed-Form SVD Alignment

Masato Kitamura and Takeshi Nishida
(The University of Kitakyushu, Japan)

This paper presents a two-stage point-cloud registration method for high-accuracy 3D object pose estimation that integrates geometric alignment with appearance cues derived from color information. Conventional distance-based registration methods, such as ICP and FGR, are sensitive to initial pose errors, partial visibility, non-uniform sampling density, and exhibit inherent ambiguity when applied to rotationally symmetric objects. In the proposed framework, an initial rigid transformation is obtained using a closed-form SVD-based alignment, which provides a stable and non-iterative geometric initialization. Subsequently, color-edge points, extracted from local variations in hue and saturation, are employed to resolve the residual in-plane rotational ambiguity through a coarse-to-fine search around the z-axis. Final refinement is performed using ICP. Simulation and real RGB-D experiments demonstrate that the proposed method substantially reduces rotational and translational errors compared with SVD and FGR, while maintaining real-time performance on standard hardware. These results indicate that combining color-derived appearance cues with geometric features is an effective mechanism for robust pose estimation under practical sensing conditions.

GS32-2 CAD-Based Stereo RGB Vehicle 6-DoF Pose Estimation under Illumination and Structural Variations

Yusuke Saito
(Independent Researcher, Japan)

Accurate six-degree-of-freedom (6-DoF) vehicle pose estimation from RGB images is key to mixed reality (MR) applications requiring alignment between digital CAD models and real-world objects, yet robust pose estimation in real environments remains challenging despite recent RGB-based methods enabling deployment on commercially available head-mounted displays. Vehicle appearance is strongly influenced by illumination conditions such as outdoor sunlight, shadows, and artificial lighting. In addition, manufacturing-related variations, including differences in component assembly and color configurations, introduce unavoidable discrepancies between pre-designed CAD models and real vehicles. These factors create a domain gap that degrades the performance and stability of existing CAD-based stereo RGB pose estimation pipelines, especially in systems without depth sensors. In this study, we analyze a CAD-based stereo RGB vehicle pose estimation framework under controlled illumination and manufacturing-related variations. We implement a multi-stage pipeline consisting of vehicle detection, pose initialization, and render-and-compare-based temporal refinement. To enable evaluation, we construct a 3D game engine-based synthetic dataset that allows control of lighting environments and structural configurations. Through

quantitative and qualitative experiments, we identify characteristic failure modes and performance degradation trends. We discuss an extension framework toward illumination-aware vehicle pose estimation and provide a foundation for future integration of advanced 3D representations in real-world MR scenarios.

GS32-3 Resolving the Sim2Real Gap through Common Style Domain Unification for Computational Object Constancy

Wataru Takahashi and Takeshi Nishida
(The University of Kitakyushu, Japan)

Addressing the Sim2Real gap, primarily caused by the "texture bias" of CNNs, is essential for utilizing synthetic data in deep learning. Conventional pixel-level adaptation methods typically require retraining for each specific target domain (e.g., sunny, rainy), limiting their real-world applicability to dynamic environments. To address this, we propose "Common Style Domain Unification," a single-source domain generalization strategy that maps both synthetic and real domains to a single, abstract common style using the SaMST style transformer. Crucially, our method generates a unified representation without using any target domain data, enabling zero-shot adaptation to unknown environments. In evaluations across diverse real-world urban driving scenarios (Cityscapes, FoggyCityscapes, RainCityscapes), our method achieved performance comparable to a target-specialized conventional method in the standard domain while outperforming it in unseen adverse weather conditions, specifically minimizing the performance drop in the foggy domain to 0.6 points compared to 3.6 points for the conventional method. These results highlight the potential of our approach for realizing computational object constancy in robust visual recognition systems.

GS32-4 Image Quality Improvement for Fukushima Daiichi Remote Operations using Denoising Prior to Super Resolution

Yuta Tanifuji, Takashi Imabuchi and Kuniaki Kawabata
(Fukushima Research and Engineering Institute, Japan Atomic Energy Agency, Japan)

Remote operation is essential for decommissioning work at the Fukushima Daiichi Nuclear Power Station (1F), where high radiation and limited accessibility restrict human intervention. Camera images acquired in underwater and in air environments are often degraded by turbidity, suspended particles, low illumination, and radiation induced sensor noise, which hinders reliable visual interpretation. In addition, clean reference images are generally unavailable, and image restoration must avoid hallucinated details from the viewpoints of safety and governance. This paper investigates a conservative image restoration pipeline suitable for such reference free conditions. We combine Noise2Noise (N2N) for denoising with FSRCNN for Super Resolution and sharpening, and adopt a two stage configuration in which denoising is applied prior to Super Resolution (N2N→FSRCNN). This ordering suppresses noise before magnification and prevents amplified artifacts. Four configurations are evaluated using real images from 1F with non reference image quality metrics and qualitative visual inspection. The results show that N2N→FSRCNN provides more stable improvements for degraded scenes, suppressing residual noise and fogging while preserving edge continuity, making it suitable for practical remote operation support at 1F.

GS32-5 Growing Neural Gas based Deep Learning Method for 3D Object Recognition

Ryo Shibata, Yuichiro Toda, Takayuki Matsuno
(Okayama University, Japan)

In recent years, autonomous robots are increasingly required to address labor shortages and operate in harsh environments. To navigate safely, such robots must analyze sensor data from real environments and generate environmental maps. Growing Neural Gas (GNG) is an unsupervised learning algorithm that dynamically constructs graph structures from 3D point clouds and is robust to downsampling and inhomogeneous data. However, while GNG enables clustering based on topological structures, it cannot directly perform object recognition. Object recognition is essential for diverse robotic tasks, and deep learning approaches have proven effective for this purpose. PointNet was the first deep learning model to handle the irregularity and permutation invariance of point clouds. In this study, we propose PointGNG, a novel model that integrates GNG with PointNet++ and extracts local structural features using adjacency-based neighborhood information. The proposed method achieves accuracy comparable to PointNet++ on the ModelNet40 benchmark and shows slight improvements on datasets containing background information. These results demonstrate that integrating unsupervised learning with deep learning via GNG enables efficient point cloud processing and robust object recognition. Furthermore, we evaluated the robustness of the proposed model and examined how the choice of training dataset affects its adaptability to real-world environments.

GS32-6 Improving Illumination Robustness in Autonomous Model Car Driving Using Deep Learning

Kosuke Nagayama, Noritaka Shigei and Yoshihiro Nakamura
(Kagoshima University, Japan)

Autonomous mobile robots operating in indoor environments often experience severe performance degradation under drastic illumination changes, particularly when training and deployment conditions differ. This study proposes a computationally efficient, real-time image preprocessing pipeline that enables illumination-invariant feature representation for stable end-to-end autonomous driving using the DonkeyCar framework. The proposed approach consists of three key steps: transformation to the YUV color space, luminance normalization using Contrast Limited Adaptive Histogram Equalization (CLAHE), and masking of unstable window regions and reflective surfaces. By applying CLAHE exclusively to the luminance channel while keeping chrominance channels fixed, variations in color temperature are suppressed without sacrificing essential structural features. In addition, masking regions prone to glare and reflection effectively reduces illumination-induced feature inversion. The proposed preprocessing pipeline was evaluated on a complex indoor course containing multiple checkpoints, including obstacle avoidance and turnaround tasks. Experimental results show that a model trained solely on daytime data achieves stable and accurate autonomous navigation under both daytime and nighttime conditions, demonstrating strong domain generalization. These findings indicate that simple, lightweight preprocessing can substantially improve robustness in vision-based indoor autonomous driving systems.

January 23 (Friday), 13:00-14:15

B1 Meeting Room 3

GS34 Sensor and multi-sensor data fusion

Chair: Yoshiaki Yamazaki (Meisei University, Japan)

GS34-1 Multi-Node eCO_2 Calibration Using Regression Models and Its Effect on Indoor Occupancy Estimation

Shijia Li, Noritaka Shigei and Yoshihiro Nakamura
(Kagoshima University, China)

Indoor environmental sensing plays a key role in energy-efficient building operation and privacy-preserving monitoring of space utilization. Estimating room occupancy from indoor CO2 concentration induced by human respiration is a widely studied non-intrusive approach; however, large-scale deployment remains challenging. Although nondispersive infrared (NDIR) CO2 sensors provide accurate measurements, their cost, size, and power consumption hinder dense installation. In contrast, compact metal-oxide (MOX) gas sensors can output estimated CO2 (eCO2) at low cost, but their readings are affected by environmental conditions such as temperature, humidity, ventilation, and background volatile organic compounds (VOCs). This study investigates whether atmospheric-pressure-aware calibration of MOX-based eCO2 enables occupancy estimation in a multi-node indoor setting. Three environmental sensor nodes were deployed at different indoor locations, and node-specific calibration models were trained using co-located MOX and NDIR measurements. The calibrated CO2 signals were then combined with illuminance, temperature, humidity, and temporal features to estimate the number of people at each location. Through systematic evaluation, we show that pressure-aware calibration stabilizes MOX-based eCO2 and often improves people-count regression relative to raw eCO2, while NDIR-based inputs still provide the most reliable performance overall.

GS34-2 Online Self-Recalibration of RGB-D Cameras in Industrial Robot Cells

Riku Kobayashi, Ukyo Taniguchi and Takeshi Nishida
(The University of Kitakyushu, Japan)

Small RGB-D cameras installed on walls or around workpieces in robot cells are often required to be flexibly repositioned and reoriented according to task requirements. However, unintended changes in camera pose caused by robot-induced vibrations or accidental contacts can lead to measurement errors, potentially resulting in task failures or degraded operational safety. This paper proposes an online method for automatically detecting pose variations of RGB-D cameras by jointly analyzing time-series signals from the built-in IMU and temporal changes in the acquired point clouds. When a pose variation is detected, the proposed system autonomously triggers a recalibration sequence using a hand–eye camera mounted on the robot to re-estimate the extrinsic parameters of the affected RGB-D camera and restore measurement accuracy. A key feature of the proposed approach is that

recalibration can be performed with minimal interruption to the ongoing robot task. Experiments conducted using an xArm6 collaborative robot and multiple Intel RealSense D435i cameras demonstrate that the proposed method can reliably detect pose variations of fixed RGB-D cameras and successfully recover measurement accuracy through autonomous recalibration in real-world environments.

GS34-3 Auto-Calibration of Multi AR-Marker-Integrated RGB-D Sensors Using Hand-Eye Camera

Ukyo Taniguchi, Riku Kobayashi and Takeshi Nishida
(The University of Kitakyushu, Japan)

To address occlusion issues inherent in robotic grasping using a single RGB-D sensor, this paper proposes a framework for integrating multiple RGB-D sensors through automated extrinsic calibration. Each sensor is equipped with an AR-marker-integrated module, and a robot-mounted hand-eye camera performs close-range observation to estimate sensor poses with high precision. By conducting calibration at close range, distance-dependent depth errors are effectively suppressed, enabling millimeter alignment accuracy. After calibration, point clouds acquired from multiple viewpoints are transformed into a unified robot coordinate system and fused to generate a complete 3D model with occluded regions reconstructed. This multi-view representation supports stable object recognition and reliable grasping performance even in complex environments.

GS34-4 Evolving the Models of Friction Coefficients of Automobile Tires via Genetic Programming

Ivan Tanev[1], Hiroshige Hasegawa[2] and Manaka Tokiwai[1]
([1]Doshisha University, Japan)
([2]Toyo Tire Corporation, Japan)

: Our research is motivated by the apparent discrepancy between (i) the importance of the real time (continuous) estimation of the maximum value of the friction coefficient (MFC) μmax between the automobile tires and the road surface and (ii) the lack of both a practically feasible and robust methods of such estimation. Our objective is investigating the possibility to develop the models of MFC as algebraic functions of various (real time) parameters, pertinent to the interaction of the tires and the road surface in a “normal”, steady state (rather than accelerating-, braking-, or turning-) driving in the four main – dry, wet, snowy, and icy – road conditions. Due to the anticipated complexity of the models of MFC, and the amount of the training set of data (timestamped values of parameters obtained during a brief driving session) it approximates, we assumed that the models cannot be handcrafted by a human expert. To develop such models, we applied genetic programming (GP) – a heuristic, automated problem-solving approach inspired by the evolution in nature. The experimentally obtained best-evolved models were acceptably accurate (with a difference between the estimated MFC and the actual one lower than 0.05) for slippery road conditions (wet, snowy, and icy).

GS34-5 Ground Sampling Distance (GSD) Metric for Vessel Motion Detection Using Cameras

Taichi Shikazono, Etsuro Shimizu and Ayako Umeda
(Tokyo University of Marine Science and Technology, Japan)

The ocean supports human life in many ways, yet maritime accidents and military threats remain major issues. In order to use the ocean effectively, ensuring the safety of maritime activities is essential, making maritime domain awareness is increasingly important. This involves observing the ocean and detecting objects at sea using satellites and aircraft. Understanding maritime activities requires effective vessel detection, and recognizing vessel motion from still images enables early responses to suspicious vessels. When performing motion recognition from above, the required sensor resolution must be defined, but the necessary GSD has not yet been clarified. This study therefore identifies the GSD needed for vessel motion recognition and proposes an indicator expressing its upper limit. This indicator allows determination of the required GSD based on target size. For variable-focus sensors, it enables appropriate adjustment, and for fixed-focus sensors, it supports requesting suitable observation platforms. It also helps define sensor performance requirements when designing future maritime monitoring technologies.

January 23 (Friday), 13:00-14:15

B1 Meeting Room 4

GS35 Walking robot & Modular robotics

Chair: Yuta Hanazawa (Kyushu Institute of Technology, Japan)

GS35-1 Learning a Unified Fault-Tolerant Gait Policy for Quadruped Robots under Single-Joint Sudden Free-Swinging Failures

Keisuke Matsuura[1] and Noritaka Sato[2]
([1]Department of Electrical and Electronic Engineering, Nagoya Institute of Technology, Japan)
([2]Department of Electrical and Mechanical Engineering, Nagoya Institute of Technology, Japan)

In recent years, quadruped robots have been expected to play an active role in urgent missions such as search and rescue operations in disaster sites. However, sudden joint failures in such harsh environments remain a significant challenge. In particular, a free-swinging joint failure (FSJF), in which the actuator completely loses torque and the joint rotates freely under external forces and gravity, is a severe failure that makes posture stabilization extremely difficult. Despite its frequent occurrence in real-world environments, most previous studies have primarily focused on locking failures, leaving FSJFs insufficiently addressed. In this study, we focus on the scenario in which “any single joint suddenly fails” in a quadruped robot and develop fault-tolerant locomotion using deep reinforcement learning. Specifically, we incorporate the failure state as an input parameter and train a single model capable of handling FSJFs across all 12 joints. Simulation experiments demonstrated that the learned policy successfully avoided falls and maintained locomotion with a high success rate, even when FSJFs occurred at random timings during walking.

GS35-2 Development of a simple learning method for bipedal robot gait control using deep reinforcement learning

Shota Yamashita[1], Hideaki Itoh[1], Hisao Fukumoto[1] and Hiroshi Wakuya[2]
([1]Electrical and Electronic Engineering Course, Graduate School of Science and Engineering, Saga University, Japan)
([2]Faculty of Education, Saga University, Japan)

In this study, a framework for enabling bipedal robots to acquire natural and stable walking behaviors through low-cost deep reinforcement learning was proposed. In the proposed approach, the walking motions generated by tracing the trajectories of a Unity-based animation model were utilized as demonstrations for Generative Adversarial Imitation Learning (GAIL), while a forward-locomotion task that prevents falling was concurrently optimized using Proximal Policy Optimization (PPO). To further enhance robustness, domain randomization, externally applied forces, and curriculum learning were incorporated so that the learned policy could be gradually adapted to increasingly challenging conditions. Through these mechanisms, stable and human-like gait patterns were successfully obtained within the simulation environment, demonstrating that the animation-driven imitation strategy can effectively reduce the cost and effort typically required for collecting real walking data. Although discrepancies were observed when the learned policy was transferred to the physical robot, the overall findings indicate that the combination of imitation learning and reinforcement learning provides a promising and practical direction for realizing affordable and efficient acquisition of bipedal locomotion skills.

GS35-3 Humanoid Robot Motion Control Using Motion Generation-Based Deep Reinforcement Learning and Large Language Models

Taichi Yoshitomi[1], Hideaki Itoh[1], Hisao Fukumoto[1] and Hiroshi Wakuya[2]
([1]Electrical and Electronic Engineering Course, Graduate School of Science and Engineering, Saga University, Japan)
([2]Faculty of Education, Saga University, Japan)

Recent advancements in deep reinforcement learning have enabled complex motion control for bipedal robots. However, designing reward functions remains challenging, often making imitation learning a preferred alternative. Nevertheless, traditional imitation learning typically requires expensive motion capture data. In addition to the high cost of data acquisition, existing systems often suffer from limited interactivity, lacking the flexibility required to interpret instructions provided in natural language. To address these dual challenges, we propose a framework that integrates generative AI into the robot training and control pipeline for both motion generation of imitation learning and high-level decision-making. Specifically, we generate expert data for imitation learning---reference motions derived from text---via a Motion Diffusion Model (MDM). Using this data, we perform training with Generative Adversarial Imitation Learning (GAIL) to mimic the expert behavior, combined with Proximal Policy Optimization (PPO) for stable policy optimization. Additionally, we integrated a Vision-Language Model (Qwen2.5-VL) to achieve target identification based on visual and linguistic information. Experimental results confirmed the functionality of the imitation learning using generated data and the associated decision-making. This study suggests the potential for realizing low-cost and linguistically interactive robot systems.

GS35-4 Improving the Efficiency of Morphology Search for Modular Robots by Forecasting Performance

Sanato Nagata and Nobutaka Kimura
(Hitachi, Ltd., Japan)

We study early prediction of final walking performance for reconfigurable legged metamachines built from autonomous modular legs. Using the public ModularLegs framework, we executed reinforcement learning (RL) for 1,029 runs with different robot morphologies and collected learning curves of episodic reward. To enable consistent modeling across runs, each curve was resampled to 100 normalized episode indices. We then predicted the final performance from an early prefix of the learning curve by featurizing the prefix and regressing the remaining improvement. We evaluated Ridge and Gradient Boosting Regression (GBR) baselines, and propose an upper-tail?aware Quantile-GBR enhanced with up-weighting for high-performing runs and isotonic calibration to reduce underestimation. Results show that prediction error increases as the observation length L becomes shorter, confirming the difficulty of early prediction. While average accuracy across all runs is comparable between GBR and Quantile-GBR, the proposed Quantile-GBR substantially reduces underestimation in the high-performance regime: the absolute error for runs whose true final performance is above the training-set mean and are underpredicted is reduced by approximately 85% compared to GBR. This risk-averse behavior is valuable for morphology search, where prematurely discarding promising designs is costly.

GS35-5 An Electromechanical Platform for Studying Pattern Formation in FloaTiles

Georgii Karelin, Brian Morrissey and Tom Froese
(Okinawa Institute of Science and Technology, Embodied Cognitive Science Unit, Japan)

This paper describes FloaTiles, an ongoing experimental project in artificial life based on inexpensive passive floating tiles that interact on a water surface through surface-tension entropic forces. The system uses simple 3D-printed elements without motors or internal control. The Cheerios effect provides the primary attractive interaction between tiles. The minimization of surface tension drives the collective motion of particles. Practical issues such as noise, humidity, evaporation, heat dissipation, and environmental airflow are also discussed. Strong perturbations break clusters apart, while weaker perturbations can allow rearrangement by releasing tiles from tightly held configurations. Tile groupings change during experiments, and repeated runs typically produce different spatial arrangements. Video is used to observe tile motion and interactions during experiments. The setup includes a shallow pool placed on a rigid table. A vibration motor with adjustable drive frequency generates water motion. Airflow and bubble injection can be added as external disturbances. All components are inexpensive and chosen for easy replacement during iterative experimentation, or if they break. Inspired by the ideas of Gilbert Simondon and Rube Goldberg, rather than producing a fixed structure, the setup remains relatively open-ended, with tile arrangements shaped continuously by material interactions rather than explicit control or design goals.

January 23 (Friday), 14:30-15:30

3F Meeting Room 33

OS9 AROB: Biomimetic Machines and Robots III

Organizers: Keigo Watanabe (IPU Tokyo, Japan / Okayama University, Japan)
Kiyotaka Izumi (Saga University, Japan)
Fusaomi Nagata (Sanyo-Onoda City University, Japan)

OS9-1 Disturbance Observer Based Rotor Wide Friction Estimation for Robust Motion Control

M.K.C. Dinesh Chinthaka[1], A.M. Harsha S. Abeykoon[2] and Lanka Udawatta[3]
([1]University of Moratuwa, Sri Lanka)
([2]Embry-Riddle Aeronautical University, United States)
([3]Higher Colleges of Technology, United Arab Emirates)

Friction is often disregarded in DC servomotor applications, especially in small motors where frictional torque is minimal compared to motor torque. However, in high-precision motion control, static and viscous friction must be accurately modeled and compensated. Traditional approaches typically estimate friction as an average value over a full rotation, which fails to capture variations when friction is position dependent. This paper proposes a refined method to estimate friction by isolating static and viscous components across 360° rotor rotation. A Disturbance Observer (DOB) is employed as a key tool in real-time torque estimation. Experimental results highlight that conventional methods underestimate static friction, and the viscous friction coefficient exhibits nonlinear behavior with varying speed. The proposed rotor-wide friction estimation approach enhances friction compensation for robust motion control and can also serve as an indicator of mechanical health, detecting issues such as rotor imbalance or bearing wear.

OS9-2 Development of a Reinforcement Learning-X-Plane Simulation Pipeline for Guidance and Control of Drone Smart Gliding Micromunition

Dynah Ruiza Opeso[1], Immanuel Paradela[2], Donato Juayang[2], Alven Bandivas[2], Maria Fe Bahinting[1], Leah Aquino-Alindayo[1] and Sherwin Guirnaldo[2]
([1]Department of Computer Applications,
Mindanao State University Iligan Institute of Technology, Philippines)
([2]Department of Mechanical Engineering and Technology,
Mindanao State University Iligan Institute of Technology, Philippines)

This paper presents a reinforcement learning (RL) and X-Plane simulation pipeline for the guidance and control of a small unpowered drone smart gliding micromunition. Unlike conventional aircraft-oriented RL–X-Plane frameworks, the proposed approach extends this integration to gliding objects characterized by strongly nonlinear dynamics and limited control authority. The system employed the Watkins Q(λ) algorithm with eligibility traces to perform adaptive gain scheduling for three independent PID controllers governing roll, pitch, and yaw axes. By training directly within X-Plane 12, the agent learns from realistic aerodynamic effects and transient flight behaviors, reducing dependence on computational fluid dynamics simulations and wind tunnel testing. Validation through 100 test cases demonstrated that the RL-tuned PID controller achieved an 87% hit rate compared to 78% for the conventional fixed-gain PID within a 10-meter impact threshold. CEP metrics showed substantial improvements, with CEP50 reduced from 3.29 m to 2.65 m and CEP90 reduced from 80.56 m to 34.75 m. These results confirm that RL-based gain scheduling effectively enhances guidance performance for gliding micromunitions, demonstrating the pipeline's viability as a practical framework for rapid prototyping and controller optimization.

OS9-3 Force-Distribution-Based Structural Analysis and Design of a Lower-Limb Prosthesis with a Variable-Stiffness Mechanism During the Gait Cycle

Kittitouch Chuvongs[1,2], Supaphon Kamon[2], Shogo Nonaka[1] and Sora Yasunobu[1]
([1]National Institute of Technology, Tsuyama College, Japan)
([2]KOSEN-KMITL, Thailand)

A prosthetic leg with a variable stiffness mechanism is an artificial limb that replaces the missing part of the leg to support the mobility of the user. With the application of a variable stiffness mechanism to improve the ankle push-off power, which is important to improve walking health. The conventional variable stiffness mechanism design is focused on only one direction, which resulted in insufficient reactive force in another direction. This inspired the design of the proposal prosthetic leg around the double spring mechanism that can react to force in every sequence of the gait cycle by using 2 compression springs with a length adjustment mechanism to change stiffness. In order to design and optimize the variable stiffness mechanism and each component that can endure the force from the user's weight and ground reaction force while minimizing the overall weight to match the weight of a normal leg as much as possible, a FEM static simulation and topology study are required. The result of the simulation allowed for optimization of the weight of the prosthetic leg and ensured that the prosthetic leg could endure the force during the gait cycle.

OS9-4 Textile-Based Surface EMG Sensors for Real-Time Driver Fatigue Detection in Intelligent Vehicle Systems

Madushi Hansika Medagedara[1] and Thrishantha Nanayakkara[2]
([1]University of Moratuwa, Sri Lanka)
([2]Imperial College, United Kingdom)

Driver fatigue is a major cause of road accidents, yet it remains difficult to detect before visible behavioral symptoms appear. This study presents a textile-based surface electromyography (sEMG) sensing approach for early identification of neuromuscular fatigue during driving. Custom knitted electrodes were fabricated using conductive yarns and optimized loop structures to ensure comfort, stability, and reliable skin contact. Sensors were integrated into a flexible seat interface and positioned over four upper-body muscles—trapezius, deltoid, rhomboid, and erector spinae—critical for posture maintenance. sEMG signals were collected at 250 Hz under simulated driving conditions and analyzed using time-, frequency-, and time–frequency-domain features such as RMS, MNF, MDF, entropy, and bispectral parameters. The results showed progressive fatigue patterns, particularly in the rhomboid and erector spinae muscles, with spectral shifts and amplitude increases consistent with fatigue progression. A lightweight rule-based decision framework was implemented for real-time interpretation. The study demonstrates the potential of textile sEMG sensors for unobtrusive, continuous monitoring in intelligent vehicle systems designed to enhance driver safety and comfort.

January 23 (Friday), 14:30-15:30

B1 Meeting Room 3

GS22 Machine learning V

Chair: TBD

GS22-1 A Project-Based Learning Approach for Teaching Deep Learning Concepts through Gesture Recognition Using Mobile Devices

Pablo Bernal Alarcón and Jae Hoon Lee
(Ehime University, Japan)

The growing demand for machine learning professionals necessitates improved educational methods, as abstract, highly mathematical concepts are challenging to teach. To enhance engagement and accessibility in higher education, project-based learning (PBL) offers a promising solution by connecting theory to practical application. This work addresses the need for more immersive learning in an introductory Convolutional Neural Networks (CNNs) course at Ehime University, which currently teaches fundamental concepts of CNNs by using the MNIST dataset. A project is proposed, which encourages students to build their own dataset, in order to, later in the course, develop and train a CNN used for motion recognition. Students use a mobile application to collect data from different gestures, which is then combined and pre-processed to be used as input data for the motion recognition CNN. This paper presents an overview of the project, which is being prepared for classroom implementation to evaluate its effectiveness in improving the learning experience for the students.

GS22-2 Evaluation of the Temporal Similarity of Clustering Results for 3D Fish Trajectories

Kohei Sugimoto and Ryosuke Saga
(Osaka Metropolitan University, Graduate School of Informatics, Japan)

In the field of aquaculture, individual fish management and anomaly detection are essential for efficient farming, and in recent years, approaches based on machine learning have been increasingly adopted instead of manual observation. However, the practicality of machine learning models in real aquaculture environments and the reusability of knowledge obtained from anomaly detection in one period for other periods have not yet been sufficiently investigated. In this study, we first estimate three-dimensional trajectories of fish schools from monocular images by combining object tracking and depth estimation, and then extract individual-level features. Based on these features, we perform clustering and examine the practicality of a method for extracting clusters regarded as anomaly. Specifically, we evaluate reusability by comparing, using Earth Mover's Distance (EMD), the similarity between the distribution at a reference time and those at the same time on different days, as well as the similarity with distributions at different times.

GS22-3 Development of a Quantitative Evaluation Method for CO_2 Absorption by Ground Cover Plants Using Sensor Data and Machine Learning

Manato Takewaki[1], Yuya Hata[2], Hiroaki Niwa[3], Tomohiro Morizane[4], Takehide Nakagawa[5], Masamoto Tafu[1] and Ryuichi Matoba[1]
([1]National Institute of Technology, Toyama, Japan)
([2]Fudo Tetra Corporation, Japan)
([3]Hokuetsu Acetylene Co., Ltd, Japan)
([4]Toyama Shoji Co., Ltd, Japan)
([5]Yatsuo Kogyo Co., Ltd, Japan)

This study aims to estimate outdoor CO_2 dynamics around groundcover plants using a low-cost environmental sensing system and machine-learning models. A measurement system was constructed using Raspberry Pi and multiple NDIR-based CO_2, temperature, humidity, and light sensors to continuously monitor environmental conditions. Time-series datasets were collected at 5-second intervals, and two models—a baseline MLP and a CNN-GRU hybrid—were evaluated. Results show that the CNN-GRU model achieved substantially higher accuracy (MAE $\approx$ 130 ppm, $R^2 \approx 0.62$) than the MLP, effectively capturing diurnal CO_2 variation. An ablation study revealed that CO_2 inflow is the dominant feature, while temperature occasionally acts as noise. This approach demonstrates the feasibility of low-cost sensing combined with deep learning for outdoor CO_2 monitoring.

GS22-4 Disentangling Action from Species: Adversarial Gated Fusion for Robust Unsupervised Animal Behavior Recognition

Mayu Kikuchi[1], Yasumasa Tamura[2] and Masahito Yamamoto[2]
([1]Graduate School of Information Science and Technology, Hokkaido University, Japan)
([2]Faculty of Information Science and Technology, Hokkaido University, Japan)

Animal behavior recognition is essential for ensuring animal welfare and effective management in zoos. However, conventional approaches require large-scale annotated datasets for every target behavior, limiting their scalability to new environments, species, and behavioral categories. To address this challenge, we extend our previous dual-stream Gated Fusion architecture by incorporating an adversarial species-disentanglement module designed to reduce species-dependent bias in the embedding space. Our framework adaptively integrates appearance cues from RGB videos and motion information from optical flow, while adversarial training encourages the embedding to become more species-invariant. We trained our model on the large-scale Animal Kingdom dataset and evaluated it on an out-of-domain polar bear surveillance dataset from Sapporo Maruyama Zoo. Experimental results show that adversarially disentangled Gated Fusion achieves clear improvements over single-modality models and provides a modest yet consistent gain compared with the non-adversarial Gated Fusion baseline in clustering metrics such as ARI and NMI. These findings indicate that introducing species-invariance can further enhance the robustness and generalization ability of multi-modal behavior embeddings in real-world zoo environments.

January 23 (Friday), 14:30-15:15

B1 Meeting Room 4

GS31 Parallel and distributed computing

Chair: Naoki Wakamiya (The University of Osaka, Japan)

GS31-1 Cooperative Sensor–Display Interfaces on a Super General-Purpose SoC

Hibiki Shinozaki and Akira Yamawaki
(Kyushu Institute of Technology, Japan)

The rapid expansion of the Internet of Things (IoT) demands diverse product solutions, yet the design costs for custom System-on-Chips (SoCs) continue to rise. To overcome this, we propose the super gEneral-Purpose sOC (EPoC) framework, which utilizes virtual hardware interfaces and dynamic reconfiguration. This approach allows a single SoC to adapt to various peripheral specifications. By combining High-Level Synthesis (HLS) with Register Transfer Level (RTL) descriptions, EPoC facilitates the rapid development of dedicated interfaces for heterogeneous peripherals. While previous work validated EPoC with simple sensors, this study extends its application to multimedia devices. We implemented hardware interfaces for an ILI9341 display and a TSC2046 touch panel, integrating them with a 9-axis sensor into a sprite-based image movement system. This cooperative processing is ideal for intuitive human–machine interfaces in robotics. Experimental evaluations on actual hardware demonstrate the framework's practicality. Specifically, the proposed method achieved up to 17 times higher energy efficiency for display rendering and approximately 8 times higher efficiency for touch panel coordinate calculations compared to software execution, proving its effectiveness for high-performance, low-power embedded systems.

GS31-2 Event-Driven Communication Cellular Automata: From Theoretical Model to Clockless Hardware Implementation (Withdrawal)

Binmao Liu[1], Jia Lee[1] and Teijiro Isokawa[2]
([1]College of Computer Science, Chongqing University, China)
([2]Graduate School of Engineering, University of Hyogo, Japan)

This paper proposes a novel cellular automaton model, the Event-Driven Communication Cellular Automaton (ECCA), based on asynchronous and event-driven communication mechanisms. By introducing event-driven updates triggered by non-trivial state transitions and incorporating an internal buffer queue in each cell, ECCA allows asynchronous reading of past states of neighboring cells while reducing redundant updates. The robustness of ECCA is analyzed through theoretical modeling, and an asynchronous circuit implementation is developed based on the proposed model. Functional simulations validate the feasibility of the hardware structure. This study provides a new framework for the physical implementation of asynchronous cellular automata.

GS31-3 An Efficient Algorithm without Data Cache to Simulate a Large Scale Flock in 3D Space

Tatsuo Unemi
(Soka University, Japan)

This paper presents an efficient algorithm to simulate flocking behavior on a compact computer, Mac mini, assuming to be portable for exhibitions. The performance operating up to one million agents is achieved through the partitioning of the 3D world space and load balancing between processing elements within the GPU and a multi-core CPU. A data cache that maintains groups of individuals for each partition or near neighbors for each individual can also effectively reduce computation costs. However, the maintenance cost of such data cache in a dynamic environment is substantial, and the memory management disrupts the effects of cache memory on the chip. To mitigate these drawbacks, the algorithm proposed here does not employ such mechanisms. Instead, it calculates the partition's position of each individual at each step. Statistical analysis of the throughput time for each phase of computation in various parameter settings indicated that such recalculations are relatively less demanding compared to the primary parallel computation responsible for the movement of individuals. In conjunction with a point cloud that represents a visitor's body shape within the virtual 3D space, the algorithm facilitated a prompt response from a large-scale flocking of agents.

GS31-4 Security Evaluation of Dataset-aware Cumulative Gradient Smash Poisoning Method in Split Federated Learning

Chisei Ishida, Ryo Kumagai, Shu Takemoto, Yusuke Nozaki and Masaya Yoshikawa
(Meijo University, Japan)

In recent years, Split Federated Learning (SFL) has gained attention as a learning framework that combines Federated Learning (FL) and Split Learning (SL). FL provides local data protection and enables parallel training, while SL enhances model confidentiality and reduces client-side computational load by splitting a model. Consequently, SFL enables efficient and secure learning. To ensure the safe deployment of SFL, it is essential to evaluate its stability from multiple perspectives. In this study, we propose Dataset-aware Cumulative Gradient Smash Poisoning (DCGSP) as a novel safety analysis method for SFL. DCGSP is a hybrid approach that combines gradient-accumulating perturbations applied to smashed data over time with perturbations in the input space. This enables the evaluation of the impact of long-term accumulated perturbations on the safety and reliability of SFL models.