# CONFERENCE PROGRAM

## 2026 11th International Conference on Frontiers of Signal Processing


## The 2026 3rd International Conference on Image Processing and Artificial Intelligence

**June 10-12, 2026 | Toulouse, France**

# Table of Content

# Welcome Message

We are pleased to welcome you to 2026 11th International Conference on Frontiers of Signal Processing (ICFSP 2026) and The 2026 3rd International Conference on Image Processing and Artificial Intelligence (IPAI 2026), held in Paris, France, during June 10-12, 2026.

ICFSP 2026 and IPAI 2026 invite authors to submit papers on different aspects of signal processing, image processing and artificial intelligence. Key areas of interest include, but are not limited to, image analysis and computational models, digital image processing and multimedia applications, next generation communication systems and signal processing.

The conferences aim to provide an interactive platform for professionals to learn about cutting-edge academic and industrial trends, share the latest scientific research and technological achievements, and discuss innovative ideas and methods. The goal is to elevate academic research and industrial application levels, supporting global initiatives around technological transformation. The conference seeks to promote technology research, development, and application both locally and internationally.

We are profoundly grateful to everyone who has helped make this event possible, including the respected authors, the invited speakers, and the peer reviewers. Special thanks also go to the conference committees for their dedication throughout the planning and execution of the conference.

We hope all participants will benefit from this event. Your contributions are essential in advancing the frontiers of knowledge and technology.

Wishing you a successful and inspiring conference experience at ICFSP 2026 and IPAI 2026!

ICFSP 2026 & IPAI 2026
Conference Committee

# Onsite Conference Information

## Conference Venue


**IRIT (Institut de Recherche en Informatique de Toulouse)**
Address: 118 Route de Narbonne, 31062 Toulouse Cedex 9, France

## Oral Presentation Tips

- ✓ The duration of a presentation slot is 15 minutes. Please prepare your presentation for about 12 minutes plus about 3 minutes for questions from the audience;
- ✓ An LCD projector and a computer will be available in every session room for regular presentations;
- ✓ Presentations MUST be uploaded at the computer at least 15 minutes before the session starts.

## Dress Code

- ✓ All participants are kindly requested to dress formally, as casual wear is discouraged.
- ✓ National formal dress is welcome.

## Attention Please

- ✓ Please ensure the safety of your belongings in public areas. For personal and property security, delegates are advised to wear their identification badges during the conference and refrain from lending them to unauthorized individuals. The conference cannot be held responsible for the loss of personal items.

**Emergency Call**: **112**

# Online Conference Information

| Zoom information |
| --- |
| **Online Room**<br><br>**ZOOM ID: 86103273769**<br><br>**ZOOM Link: https://us02web.zoom.us/j/86103273769**<br><br>**Password: 061012** |

**Time Zone**

- ✓ Toulouse Local Time: UTC+2 (CEST)
- ✓ Please make sure that both the clock and the time zone on your computer are set to the correct Paris standard time.

**Sign in and Join**

- ✓ Join a meeting without signing in: A Zoom account is not required if you join a meeting as a participant, but you cannot change the virtual background or edit the profile picture.
- ✓ Sign in with a Zoom account: All the functions are available.

**Additional Suggestions**

- ✓ A computer with an internet connection (wired connection recommended)
- ✓ USB plug-in headset with a microphone (recommended for optimal audio quality)
- ✓ Webcam (optional): built-in or USB plug-in
- ✓ Stable Internet Connection
- ✓ Quiet environment
- ✓ Proper lighting

**Presentation Tips**

Each presentation slot is 15 minutes. Please prepare to speak for around 12 minutes, allowing 3 minutes for audience questions.
Join the meeting room at least 15 minutes before the session begins.

# June 10, 2026 | Wednesday
# Toulouse Local Time: UTC+2 (CEST)

| For Onsite Participants | |
|---|---|
| **Venue:The Entrance of IRIT<br>118 Route de Narbonne, 31062 Toulouse Cedex 9, France** | |
| 14:00-17:00 | Sign in and Collect Conference Materials |

# June 10, 2026 | Wednesday
# Toulouse Local Time: UTC+2 (CEST)

| For Online Participants | |
|---|---|
| **ZOOM ID: 86103273769<br>ZOOM Link: https://us02web.zoom.us/j/86103273769<br>Passcode: 061012** | |
| 10:00-12:00 | ZOOM Test for Online Participants |

# Daily Schedule

## June 11, 2026 | Thursday
## Toulouse Local Time: UTC+2 (CEST)

**Venue: The Auditorium (in front of the building entrance) IRIT**

| | |
|---|---|
| 09:00-09:30 | **Sign In & Morning Reception** |
| **Opening Address & Keynote Speech** | |
| 09:30-09:35 | **Conference Chair Opening Address**<br>**Jacques Blanc-Talon**, Université Paris XI, France |
| 09:35-10:15 | ***Keynote Speech I: Point Cloud Coding and Quality Assessment: challenges and perspectives***<br>**Frederic Dufaux,** CNRS · Laboratoire des Signaux et Systèmes, France, IEEE Fellow |
| 10:15-10:40 | **Break Time & Group Photo** |
| 10:40-11:20 | ***Keynote Speech II: TBA***<br>**Ali Mohammad-Djafari,** Former Research Director from CNRS, France<br>International Science Consulting & Training (ISCT)<br>Zhejiang Shangfeng Company Scientific Director |
| 11:20-11:50 | ***Invited Speech I: Uplink Sum-Rate Optimization in RIS-Assisted NOMA Systems with Spatial Correlation***<br>**Luca Pallotta,** University of Basilicata, Italy |
| 12:00-14:00 | **Lunch \| The Hall of IRIT** |

| Author Presentation Sessions | | |
|---|---|---|
| **Venue: IRIT** | | |
| 14:00-15:45 | **Onsite Session 1**<br>**Biomedical Signal Processing and Health Informatics**<br><br>Session Chair: Iker Gondra, St. Francis Xavier University, Canada<br><br>FP1008, FP1005, FP2015, FP2013, FP3020, FP608, FP2018 | **The Defence room (salle des thèses)** |
| 15:45-16:10 | **Break Time** | |
| 16:10-17:55 | **Onsite Session 2**<br>**Intelligent Image Processing and Computational Vision**<br><br>Session Chair: Adriano de Oliveira Andrade, Federal University of Uberlândia, Brazil<br><br>FP1002, FP2016, FP3022-A, FP3024, FP4042, FP4041, FP506-A | **The Defence room (salle des thèses)** |
| 19:00-20:30 | **Dinner \| The Hall of IRIT** | |

# Daily Schedule

## June 12, 2026 | Friday
## Toulouse Local Time: UTC+2 (CEST)

| Online Presentation Session | | |
|---|---|---|
| 10:00-11:00 | **Online Session**<br>**Cross-Modal Intelligence and Emerging Applications**<br><br>Session Chair: Paulo Batista, University of Évora, Portugal<br><br>Invited Speaker-Paulo Batista, Invited Speaker-Ali H. Wheeb, Invited Speaker-P. Aruna, | **Online Room**<br>**ZOOM ID:**<br>**86103273769**<br>**Password: 061012** |
| Break for 10 minutes | | |
| 11:10-12:25 | **Online Session**<br>**Cross-Modal Intelligence and Emerging Applications**<br><br>Session Chair: Paulo Batista, University of Évora, Portugal<br><br>FP1007, FP4035, FP502, FP711, FP3021 | **Online Room**<br>**ZOOM ID:**<br>**86103273769**<br>**Password: 061012** |
| 12:25 | **Break Time** | |

# Keynote Speaker

**Frederic Dufaux**

**CNRS · Laboratoire des Signaux et Systèmes, France**
**IEEE Fellow**

**Speech Time: 09:35-10:15**
**Venue: The Auditorium (in front of the building entrance) IRIT**

## *Speech Title: Point Cloud Coding and Quality Assessment: challenges and perspectives*

**Abstract:** Nowadays, thanks to rapid technological progresses over the last decades, digital images and video sequences are ubiquitous, with many remarkable and successful applications and services. A key driver to research and development activities has been the objective to provide an ever-improving user experience. In particular, immersive applications aim at giving to the user the sense of being present and immersed in one location or environment, without being physically there.
In this context, a point cloud is a promising representation. A point cloud is a set of points in the 3D space represented by their spatial coordinates and additional attributes, including color pixel values, normals, or reflectance. They are often very large, easily ranging in the millions of points, and are typically sparse. Besides, one major distinguishing feature of point clouds is that, unlike images, they do not have a regular structure. For these reasons, point cloud processing presents significant challenges. In this talk, I will focus on point cloud coding and quality assessment. More specifically, I will present some recent learning-based approaches for point cloud compression and quality assessment.

**Bio:** Dr. Frederic Dufaux is a CNRS Research Director at Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes (L2S, UMR 8506), where he is head of the Telecom and Networking research hub. He is a Fellow of IEEE. Frederic received the M.Sc. in physics and Ph.D. in electrical engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1990 and 1994 respectively. He has over 30 years of experience in research, previously holding positions at EPFL, Emitall Surveillance, Genimedia, Compaq, Digital Equipment, and MIT. Frederic was Vice General Chair of ICIP 2014, General Chair of MMSP 2018, and Technical Program co-Chair of ICIP 2019 and ICIP 2021. He is Technical Program co-Chair of ICIP 2025 and MMSP 2025, and General Chair of ICME 2026. He served as Chair of the IEEE SPS Multimedia Signal Processing (MMSP) Technical Committee in 2018 and 2019. He was a member of the IEEE SPS Technical Directions Board from 2018 to 2021. He was Chair of the Steering Committee of ICME in 2022 and 2023. Since 2025, he is IEEE SPS Vice President Technical Directions, and member of the IEEE SPS Board of Governors and Executive Committee. He was also a founding member and the Chair of the EURASIP Technical Area Committee on Visual Information Processing from 2015 to 2021. He was Editor-in-Chief of Signal Processing: Image Communication from 2010 until 2019. Since 2021, he is Specialty Chief Editor of the section on Image Processing in the journal Frontiers in Signal Processing. In 2022, he received the EURASIP Meritorious Service Award, “for his leadership and contributions for the development of visual information processing within EURASIP.” Frederic is on the Executive Board of Systematic Paris-Region since 2019, a European competitiveness cluster which brings together and drives an ecosystem of excellence in digital

technologies and DeepTech. He has been involved in the standardization of digital video and imaging technologies for more than 15 years, participating both in the MPEG and JPEG committees. He was co-chairman of JPEG 2000 over wireless (JPWL) and co-chairman of JPSearch. He is the recipient of two ISO awards for these contributions. His research interests include image and video coding, 3D video, high dynamic range imaging, visual quality assessment, video surveillance, privacy protection, image and video analysis, multimedia content search and retrieval, video transmission over wireless network. He is author or co-author of 3 books, more than 250 research publications (h-index=53, 11000+ citations) and more than 25 patents issued or pending. He is in the « World′ s Top 2% Scientists » list from Stanford University.

# Keynote Speaker

**Ali Mohammad-Djafari**

**Former Research Director from CNRS, France**
**International Science Consulting & Training (ISCT)**
**Zhejiang Shangfeng Company Scientific Director**

**Speech Time:10:40-11:20(Online)**
**Venue: The Auditorium**
**(in front of the building entrance) IRIT**

## *Speech Title: TBA*

**Abstract:** TBA

**Bio:** Ali Mohammad-Djafari received the B.Sc. in electrical engineering from Polytechnic of Teheran, in 1975, the M.Sc. from Supélec (Now CentraleSupélec) in 1977, the "Docteur-Ingénieur" (Ph.D.) and "Doctorat d'Etat" in Physics, from the University of Paris Sud 11 (UPS), Orsay, France, respectively in 1981 and 1987.
He has been Research Director at CNRS, and Professor of universities in France and in many other countries until his retirement in 2018. Now, part-time, in China as the Chief Scientist in Shanfeng company, Shaoxing, Zhejiang Province. His main scientific interests include: - Developing new methods based on Bayesian inference and Information Theory approaches for Inverse Problems, and in general, in all aspects of data processing, and more specifically in imaging and vision systems. Multivariate and multi dimensional data, space-time signal and image processing, data mining, clustering, classification, machine learning and artificial intelligence methods for Diagnostics and preventive maintenance. The main application domains of his interests are Medical or biological imaging, Computed Tomography, Non Destructive Testing (NDT), fault diagnostic and preventive maintenance in industry. He has supervised more than 22 Ph.D., more than 20 Post-doc research activities and more than 50 M.Sc. Student research projects. He has more than 400 papers in journals, national and international conferences. He has organized more than 10 international workshops and conferences in France and international. He has been expert for a great number of French national and international research projects. He also participated and managed many industrial contracts with many French national industries such as EDF, RENAULT, THALES, SAFRAN and great research institutions such as CEA, INSERM, INRIA as well as the regional (such as Digiteo), national (such as ANR) and European projects (such as ERASYSBIO). In China since 2020 as a scientific talent, he obtained many scientific and technological development prizes in Zhejiang province.

# Invited Speaker

**Luca Pallotta,**

**University of Basilicata, Italy**

**Speech Time: 11:20-11:50**
**Venue: The Auditorium**
**(in front of the building entrance) IRIT**

## *Speech Title: Uplink Sum-Rate Optimization in RIS-Assisted NOMA Systems with Spatial Correlation*

**Abstarct:** This paper addresses the uplink sum-rate maximization in RIS-assisted power-domain NOMA systems. The aim is to optimize RIS phase shifts under unit-modulus constraints, resulting in a non-convex problem with coupled variables. To tackle this, an alternating optimization (AO) method is proposed, in which the RIS configuration is updated for one user at a time, keeping others fixed. Each step uses a phase-only conjugate gradient method (CGM), inspired by adaptive array processing. Various models for spatial correlation among the RIS elements are explicitly considered. Simulations show that the proposed AO-CGM approach improves both the sum-rate and the user fairness compared to conventional methods.

**Bio:** Luca Pallotta (Senior Member IEEE) received the Laurea Specialistica degree (cum laude) in telecommunication engineering in 2009 from the University of Sannio, Benevento, Italy, and the Ph.D. degree in electronic and telecommunication engineering in 2014 from the University of Naples Federico II, Naples, Italy. He was an Assistant Professor at University of Roma Tre from 2019 to 2022. He is currently an Associate Professor at University of Basilicata. More than 15 years of experience in research in Signal Processing and Remote Sensing, with strong expertise in the field of covariance matrix estimation for radar applications, radar targets detection, automatic target recognition (ATR), micro-Doppler signature analysis and exploitation, multi-polarimetric SAR image classification, statistical signal processing with emphasis on radar/SAR signal processing. Since November 2020 he is Associate Editor for IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (JSTARS). From July 2018 to February 2021, he was Associate Editor for the journal Springer Signal, Image and Video Processing (SIVP). He was a finalist of the M. Barry Carlton Award with the paper "A Novel Algorithm for Radar Classification based on Doppler Characteristics Exploiting Orthogonal Pseudo-Zernike Polynomials by C. Clemente, L. Pallotta, A. De Maio, J. J. Soraghan and A. Farina", published on IEEE Trans. on Aerospace and Electronic Systems in 2015. He won the Student Paper Competition at the IEEE Radar Conference 2013. Reviewer for several scientific international journals and conferences. Extensive publication output with more than 50 Journal papers (IEEE, IET), and more than 40 conference proceedings in highly rated conferences (IEEE Radar Conference, IRS, Nato Specialists Meeting, ...), and 7 book chapters.

# Invited Speaker

**Paulo Batista**
**University of Évora, Portugal**

**Speech Time: 10:00-10:20, June 12, 2026 | Friday**
**Online Session**

## *Speech Title: Impact of Information on the Enterprise and the Individual*

**Abstract:** Following the Second World War an explosion in the quantity of documentation led to a dramatic change in Archiving, or the profession referred to as records managers/records management and archivists/archives. Starting in the 1980s, however, archivists in Quebec began to make great progress by changing their approach and looking at the entire documentary cycle from current to definitive information. Carol Couture and Jean- Yves Rousseau made a crucial contribution towards the understanding of the Three Age Theory that viewed Archiving as an integrated discipline centered on a structural understanding of archives. In 1994, their work Les Fondements de la Discipline Archivistique, presented a new interpretation of Theodore Schellenberg's Three Age Theory. They called attention to the fact that the three phases of archival documents are not separate but, on the contrary, integrated. They argued that these three stages can even be looked at in a segmented way, provided the union between them is ensured. Their great innovation relative to Schellenberg's work lay, precisely, in critiquing the division and separation between the three ages of archival documents. Couture and Rousseau thereby brought together all the phases of the lifecycle of records, from production to dissemination, in opposition to the sterile distinction advocated by traditional archivists and document managers. In my opinion, however, the best approach to integrating information management is known as records continuum, which place archives in a post-custodial, informational, and scientific paradigm. This Australian concept arose in the 1990s amid the huge explosion of information, communication technologies and new media. This context forced Information Science to redefine its object of study. Records continuum is closely related to the integrated management model of Couture and Rousseau, while it carries their innovation further, perfecting it and replacing it with systemic dynamics and providing continuity between archives. In fact, records continuum means, literally, continuous management. It looks at the whole process from the production of records to their final archiving. Otherwise, we cannot speak of continuous management. That is why, when we speak of rigid archives – current, intermediate, and definitive, this approach is more theoretical than practical. There is, in fact, no separation between these phases, even less so from the point of view of the value of documents. The traditional distinction between information with probative and historical value ceases to exist. The information is simultaneous and is, in fact, the same.

**Bio:** Current director of the Arquivo Nacional Torre do Tombo, he was senior technician positions at the Instituto de Arquivos Nacionais/Torre do Tombo, Instituto Português do Património Cultural and the Instituto Português do Património Arquitetónico. He has also worked as researcher at the Instituto de Investigação Científica

Tropical – Centro de Estudos de História e Cartografia Antiga, and as professor at the MS program in Information Science and Documentation at Universidade NOVA de Lisboa (UNL).

Paulo Batista is PhD Researcher at CIDEHUS.UÉ-Interdisciplinary Center for History, Cultures and Societies of the University of Évora, Portugal, where is the coordinator of the research group 2: Heritage and Literacies, and professor at the Autonomous University of Lisbon, where is coordinator and professor of the Postgraduate in Promotion and Cultural and Educational Dynamization of Archives and Libraries, and the Postgraduate in Architectural Archives. Paulo Batista holds a Ph.D. in Documentation (University of Alcalá, Madrid-UAH), an MS in Information Science and Documentation - Archival Studies (UNL), and an MA in Documentation (UAH). As part of his doctorate, he also received a Diploma of Advanced Studies in Bibliography and Documentation Retrospective in Humanities (UAH), and he also holds a postgraduate degree in Information Society Law (University of Lisbon) and Information and Documentation Science - Librarianship and Archival Studies (UNL), and a specialization in Good Practices in Patrimonial Management (UNL) and Information Science and Documentation - Archival Studies (UNL). He holds an undergraduate degree in History (University of Lisbon). Paulo Batista is the author of several books and about 90 papers published in international journals and conference proceedings. He was also keynote speaker and invited speaker at various international conferences (Portugal, Argentina, Belgium, Brazil, China, Ecuador, Egypt, England, Fiji, France, India, South Africa, Thailand, Türkiye and South Korea). More informations: https://www.cienciavitae.pt//0618-CE7B-71

# Invited Speaker

**Ali H. Wheeb**
**University of Baghdad, Baghdad**

**Speech Time: 10:20-10:40, June 12, 2026 | Friday**
**Online Session**

## *Speech Title:Intelligence in the Sky: Reinforcement Learning-Driven Routing for Autonomous UAV Networks*

**Abstract:** The deployment of Unmanned Aerial Vehicle (UAV) networks and Flying Ad-Hoc Networks (FANETs) is critical for time-sensitive, high-mobility applications such as disaster response, environmental monitoring, and autonomous surveillance. However, the operational reality of these networks presents severe challenges: highly dynamic topologies, frequent link breakages, and limited power constraints. Conventionally, static routing protocols are increasingly inadequate for handling such unpredictable flying environments. This presentation introduces a cutting-edge paradigm that leverages Artificial Intelligence (AI) to achieve resilient, autonomous flying communications. Explore the integration of Reinforcement Learning (RL) into routing protocol design, allowing UAV nodes to dynamically learn optimal data paths and adapt to real-time topology changes on the fly. Furthermore, bridge the gap between abstract network algorithms and physical flight constraints by incorporating realistic, aerodynamically bounded mobility models such as 3DGM into the design process. By combining AI-driven adaptability with physical mobility constraints, this approach significantly enhances throughput, reduces latency, and optimizes energy consumption across the swarm. The speech will detail the architecture of these novel RL-routing mechanisms, explain the integration of simulation tool and AI tool to evaluate the performance, and discuss the future of fully autonomous, self-healing aerial networks.

**Bio:** Dr. Ali H. Wheeb is an Associate Professor at the College of Engineering, University of Baghdad, Iraq. He received a Ph.D. Degree with Award Graduate on Time (Ph.D. with GoT) in Electrical and Electronic Engineering from the University of UKM, Malaysia. His research interests include UAV Networks, UAV, Computer Networks, IoT, MANET, FANET, Wireless Communication, Routing protocol, Machine learning (ML), Reinforcement Learning (RL), Q-learning, Deep Q-Network (DQN), Mobility Models, and Simulation (NS-2, NS-3, ns-gym). Asst. Prof. Dr. Ali authored 22 research papers and books. Further, he has been an expert reviewer in 110 journals and conferences and reviewed more than 835 papers. Further, he was appointed as Chair of the organization at the DECA 2022 international conference and Chair of publication at the ICIDIT 2023 and ICBIS 2024 international conferences. Asst. Prof. Dr. Ali was selected as a program committee member (PCM) and technical committee member (TCM) at +50 international conferences. Further, Prof. Ali was appointed Associated Editor and Editorial Board Member in several international journals. Moreover, Asst.

Prof. Ali has received several Awards including the "Young Scientist Award" in the International Scientist Awards 2022 on Engineering, Science, and Medicine from INSO AWARDS, an Excellent Reviewer award from WoS in 2023, and GoT Award from UKM in 2024. In Addition, Asst. Prof Dr. Ali is a member of several organizations such as the Iraqi Engineering Union, European Alliance for Innovation (EAI), Asian Council of Science Editors (ACSE), Time Higher Education (THE), Franklin London Journals Press, Elsevier, IGI, Springer Nature, and STM.

# Invited Speaker

**P. Aruna**
**Q. Coimbatore Institute of Technology, India**

**Speech Time: 10:40-11:00, June 12, 2026 | Friday**
**Online Session**

## *Speech Title:Big Data Analytics in Cloud Computing*

**Abstract:** Big data analytics in the cloud offers scalability, cost-efficiency, speed, and agility. It empowers organizations with advanced analytics capabilities, global accessibility, robust security, and compliance. Embracing cloud-based analytics enables businesses to harness the full potential of their data for strategic decision-making and innovation.

**Bio:**Dr. P. Aruna is a dedicated academician with over two decades of teaching experience in Computer Science, who has consistently fostered students' academic development through effective teaching, mentoring, and active participation in institutional development. She earned her Ph.D. in 2017, specializing in cloud computing with a focus on virtual machine scheduling in cloud datacenters, aiming to optimize resource allocation, enhance resource utilization, and build energy-efficient cloud environments. With a robust research background, she has published numerous papers in reputable Scopus and Web of Science indexed journals, and her research interests span cloud computing, resource management, energy efficiency, and the integration of machine learning techniques in cloud environments. Additionally, as the Faculty Advisor of the FOSS-CIT Club dedicated to promoting open-source technologies among students, she has led the club in organizing seminars, coding competitions, and hackathons to cultivate students' technical competencies and engagement with open-source communities. Committed to establishing a vibrant research and innovation ecosystem at her institution, she motivates students to engage in research activities, obtain professional certifications, and pursue lifelong learning, with the vision of bridging the academia-industry divide and nurturing technical excellence among students.

# Onsite Session 1

## Biomedical Signal Processing and Health Informatics

- **Session Chair:** Iker Gondra, St. Francis Xavier University, Canada
- **Time:** 14:00-15:45, June 11, 2026[Toulouse Local Time: UTC+2 (CEST)]
- **Room:** The Defence room (salle des thèses), IRIT
- **Papers:** FP1008, FP1005, FP2015, FP3020, FP2013, FP608, FP2018

| | |
|---|---|
| FP1008<br>14:00-14:15 | **Sound Event Candidate Generation by Rule-Based Signal Processing**<br><br>Author(s): Leon Grude, Jessica Raabe, Gurubaran Raveendran, Thomas Bartels, Helen Schomburg, Kai Homeyer, Martin Streitenberger, Jörn Ostermann<br>Presenter: Leon Grude, University of Applied Sciences and Arts Hannover, Germany<br><br>Abstract: Automated sound event detection (SED) in long-du-ration recordings often relies on supervised learning methods that require large amounts of annotated training data. Candidate gen-eration using rule-based signal processing can reduce annotation effort by condensing continuous audio into temporally localised event candidates for targeted verification. While such approaches are widely used in wild-bird bioacoustics, their suitability for live-stock barn audio has not been systematically evaluated. This study addresses this gap by systematically comparing three rule-based detection methods—Fixed-Band Energy Detection (FBED), Time–Frequency Template-Matching Detection (TFTMD), and Adap-tive Band-Selection Energy Detection (ABSED)—for automated candidate generation in turkey barn audio. The methods are eval-uated within a unified candidate-extraction and evaluation pipe-line that standardises all processing stages except the method-spe-cific score computation. Experiments use annotated recordings of turkey vocalisation events (gobble chorus bursts) acquired under realistic barn conditions. All methods achieve high detection per-formance for the considered event type. TFTMD provides the most favourable trade-off for candidate generation in annotation workflows, achieving the highest precision and F1-score and the lowest screening overhead (review effort per confirmed event). FBED achieves the highest recall at the cost of more false positives, while ABSED provides an intermediate trade-off. These results demonstrate that rule-based signal-processing candidate genera-tion can provide an effective component for annotation workflows in SED pipelines operating under realistic barn conditions, partic-ularly for acoustically salient vocalisation events. |
| FP1005<br>14:15-14:30 | **Real-time quality-aware embedded signal processing for respiratory and cardiac peak detection in bioimpedance**<br><br>Author(s): Toms Račinskis, Armands Ancāns, Sandis Kaģis, Klara Kuzņecova, |

| | |
|---|---|
| | Modris Greitāns<br>Presenter: Toms Račinskis, Institute of Electronics and Computer Science, Latvia<br><br>**Abstract:** Driven by an aging population and medical staff shortages, continuous remote patient monitoring has become a clinical necessity. While wearable health devices are proliferating, their clinical adoption remains limited by the high power burden of processing continuous raw data and a lack of trust in algorithm outputs corrupted by real-world motion artifacts. Bioimpedance (BI) is highly promising for addressing the hardware constraints of ultra-low-power wearables, enabling simultaneous extraction of heart and respiratory rates from a single sensor channel. However, ensuring reliable, real-time on-device extraction is challenging. In response to this, we propose a signal processing method utilizing a four-stage pipeline: linear detrending, bandpass filtering, the Smoothed Recursive Moving Average Crossover (SRMAC) algorithm for peak detection, and a Signal Quality Index (SQI) based on peak morphology correlation. The algorithm was implemented in C and deployed on an ESP32-C3-Mini. Validation was performed using data collected from four subjects during resting and physical activity, comparing the results with a reference acquired from commercial sensors. The system achieved an average Root Mean Square Error (RMSE) of 4.72 beats per minute (BPM) for heart rate and 0.96 breaths per minute (BrPM) for respiratory rate. After removing signal segments marked as low quality using the SQI, the heart rate RMSE dropped to 2.66 BPM, and the respiratory rate RMSE decreased to 0.805 BrPM. Computational benchmarking demonstrated very high efficiency, requiring only approximately 2 ms of CPU time per second of processed data and utilizing only 3.5 kilobytes (KB) of Flash and 8.9 KB of DRAM per block instance. These results confirm the viability of the proposed block for low-power, long-term wearable monitoring applications. |
| FP2015<br>14:30-14:45 | **Uncovering hand tremor patterns in individuals with Parkinson's disease using Matrix Profile**<br><br>Author(s): Ariana Moura Cabral, Nicoly Augusta S. Quezada Santos, Bruno Lima Pessôa, Clarissa Cardoso dos Santos Couto Paz, Eduardo Rocon, Adriano de Oliveira Andrade<br>Presenter: Adriano de Oliveira Andrade, Federal University of Uberlândia, Brazil<br><br>**Abstract:** The analysis of involuntary movements is crucial for comprehending motor patterns in movement disorders, including Parkinson’s disease (PD). A significant challenge in identifying these patterns is the intrinsic variability found in time-series data obtained from movement detection sensors (e.g., inertial sensors). This research presents the application of Matrix Profile, a technique for automated pattern recognition in time-series data, to identify involuntary motor patterns in individuals with PD. We evaluated the method using a dataset comprising ten individuals during the ON and OFF phases of medication at rest and action conditions. The results highlight the existence of typical recurrent patterns throughout the entire duration of the time-series, specially while the participant was |

| | |
|---|---|
| | at rest. The automatic detection of patterns in PD can contribute to the characterization of the typical motor condition of patients, which can help health professionals to provide more effective treatment to patients. |
| FP3020<br>14:45-15:00 | **Interpretable Compression of Multivariate Time Series via Geometric Volume Compression**<br><br>Author(s): S. B. Belhaouari , K. Mohammadi , I. Belhaouari , Y. C. Kahalan<br>Presenter: Samir Brahim Belhaouari, Hamad Bin Khalifa University, Qatar<br><br>**Abstract:** We present a framework that is both interpretable and parameter efficient for the compression of multivariate time series (MTS) based on the Geometric Path Features (GPF). The core contribution is a two stage pipeline with (i) a novel channel selection technique ranking channels based on their average distance to all others where here we use a Shape Based Distance (SBD), which keeps only the most geometrically distinct channels and (ii) which is a grouped k-volume compression scheme collapsing the selected channels into a smaller subset of series with each capturing the signed or unsigned geometric volume swept by the group of channels over time. Path signatures are computed on these compressed series in order to retrieve compact feature vectors. On a 963-channel traffic dataset, our full pipeline compresses to 60 channels while achieving 81.2% accuracy which is 98.8% of the full signature baseline accuracy while notably using 64× fewer features. On a 61-channel medical sensing benchmark, the bivariate compressed variant outperforms the baseline (72.2% vs. 70.7%) with 61× fewer features. Empirically, the k-volume compression is robust to channel selection strategy where random selection matches SBD in accuracy which confirms that the grouped geometric compression itself is the key contribution. |
| FP2013<br>15:00-15:15 | **Accelerometer-Based Signal Processing for Objective Assessment of Wrist Rigidity in Parkinson's Disease**<br><br>Author(s): Camille Marques Alves, Yann Morere and Adriano de Oliveira Andrade<br>Presenter: Adriano de Oliveira Andrade, Federal University of Uberlândia, Brazil<br><br>**Abstract:** Parkinson’s disease is a neurodegenerative disorder associated with several motor impairments, including rigidity, which is typically evaluated through subjective clinical examination. The development of objective signal based methods for rigidity assessment has therefore become an important research topic in biomedical signal processing. This work presents an accelerometer based signal processing framework for the quantitative evaluation of wrist rigidity during controlled motor tasks. Wrist flexion and extension movements were recorded using wearable inertial sensors during active movements, slow passive movements, and fast passive movements. The acquired accelerometer signals were processed through a signal processing pipeline including filtering, envelope estimation, and spectral analysis. From these signals, motion related features were extracted in both time and |

| | |
|---|---|
| | frequency domains in order to characterize movement dynamics. The extracted features were used to compare movement patterns between individuals with Parkinson disease and healthy subjects. The results show that accelerometer derived features provide relevant information about movement dynamics and allow differentiation between groups. In particular, movements performed at higher passive velocities produced clearer differences in signal characteristics associated with rigidity. These findings indicate that accelerometer based signal processing methods can support the development of objective tools for the quantitative assessment of rigidity and may contribute to improved monitoring of motor symptoms in Parkinson disease. |
| FP608<br>15:15-15:30 | **Upscaling Human Face Geometry Images Through Super-Resolution**<br><br>Author(s): Yusuf Kızılkaya, Uluğ Bayazıt<br>Presenter: Yusuf Kızılkaya, Istanbul Technical University, Türkiye<br><br>**Abstract:** High quality 3D human face meshes are at a high demand in many applications of computer graphics as well as scientific research in general, yet high quality meshes are hard to acquire, create and store. In this paper we propose a geometry image based method that can upscale a lower quality 3D mesh face data using geometry image super-resolution. Our approach is to convert a 3D mesh into a geometry image, apply image super resolution and then convert the resulting image back into a 3D mesh. This final reconstruction is done smartly using a greedy connectivity finder algorithm. Our approach is significantly faster than existing methods of mesh refinement. This approach retains the 3D mesh structure, no conversion from a voxel model or point cloud is needed. It is also greatly flexible, working with irregular meshes and meshes containing holes. Final reconstruction step can also be swapped out for a simple regular triangulation step to receive a regular mesh, even if the original mesh was irregular. Our approach integrates existing and well proven image super resolution approaches, meaning that future improvements in image super-resolution could potentially be combined with this approach to yield better results. |
| FP2018<br>15:30-15:45 | **Efficient Real-Time Facial Emotion Recognition via Optimized ResNet18 Feature Extraction**<br><br>Author(s): Hamza Grain, Youssef Hayani, Yanis Blot--Elmazouzi, Amadou Diallo, Hamza Bayd, Imene Sekkiou, Baptiste Magnier<br>Presenter: Baptiste Magnier, IMT Mines Ales, France<br><br>**Abstract:** Facial emotion recognition constitutes a central challenge in computer vision, where the objective is to infer affective states from subtle and often ambiguous facial patterns under real world conditions. In this work, we develop a deep learning–based framework capable of performing real time emotion classification while preserving high predictive accuracy and computational |

| | |
|---|---|
| | efficiency. Building on a detailed analysis of state of the art facial emotion recognition architectures, we implement a fine tuned ResNet18 model optimized for low latency inference and robust feature extraction. This configuration achieves an effective compromise between representational capacity and model complexity, enabling deployment on resource constrained platforms. The resulting system demonstrates strong applicability across diverse vision centric domains including embedded perception modules, clinical monitoring pipelines, and adaptive learning environments highlighting the operational relevance and scalability of modern emotion recognition approaches. |

# Onsite Session 2

## Intelligent Image Processing and Computational Vision

- **Session Chair:** Adriano de Oliveira Andrade, Federal University of Uberlândia, Brazil
- **Time:** 16:10-17:55, June 11, 2026 [Toulouse Local Time: UTC+2 (CEST)]
- **Room:** The Defence room (salle des thèses), IRIT
- **Papers:** FP1002, FP2016, FP3022-A, FP3024, FP4042, FP4041, FP506-A

| | |
|---|---|
| FP1002<br>16:10-16:25 | **Learning-Free Image Similarity via Patch-Based Normalized Information Distance**<br><br>Author(s): Iker Gondra<br>Presenter: Iker Gondra, St. Francis Xavier University, Canada<br><br>**Abstract:** The Normalized Information Distance (NID), rooted in Kolmogorov complexity, provides a universal, feature-independent notion of similarity. In earlier work, we investigated the application of compression-based approximations of the NID to content-based image retrieval, demonstrating that a statistically meaningful image similarity measure could be obtained directly from raw pixel data without explicit feature extraction. However, while the conceptual soundness of the approach was validated, the method relied on whole-image compression, which limits sensitivity to localized visual structure and is poorly matched to the finite context of practical compressors. In this paper, we revisit and extend that framework by reformulating this compression-based approximation of the NID at the level of local image patches. Rather than treating images as monolithic byte streams, the proposed method decomposes images into compact, spatially aligned regions, allowing the compressor to operate within its effective context window. We evaluate the proposed approach on two benchmark datasets with increasing visual variability. Experimental results show that the proposed approach achieves particularly strong nearest-neighbor retrieval performance when there is limited intra-class variability, while performance degrades as variability increases, consistent with the expected behavior of compression-based, non-learned similarity measures. The patch-based formulation improves retrieval accuracy over both a global compression baseline and a grayscale histogram baseline. |
| FP2016<br>16:25-16:40 | **Semi-Supervised Auto-White Balancing Method for RGB-IR Images based on IR-Guidance**<br><br>Author(s): Katarina Tolja, Donik Vršnak, Marko Subašić, Sven Lončarić<br>Presenter: Katarina Tolja, University of Zagreb, Croatia<br><br>**Abstract:** Auto white balance (AWB) is a key stage in the signal processing (ISP) pipeline that automatically adjusts the colour temperature of an image to remove |

| | |
|---|---|
| | colour casts so that white or neutral objects appear indifferent under varying light sources. While existing AWB methods are developed for classic RGB cameras, emerging RGB-IR sensors introduce additional distortions due to infrared leakage and spectral cross-talk. In this work, we address white balancing for RGB-IR imaging pipelines following the learned IR leakage correction stage. We propose an IR-guided white balance network that predicts spatial gain maps using both IR-free RGB signals and the infrared channel. The IR-guided white balance module is trained on controlled laboratory conditions and refined on a real-world dataset. To achieve the domain shift, we further propose a semi-supervised Mean-Teacher training framework that leverages unlabeled real-world RGB-IR captures through consistency regularisation. The model is trained using supervised neutrality constraints from the ColorChecker last row patches, captured under controlled illumination and refined using unlabeled scenes with unknown and mixed lighting conditions. Experiments demonstrate that the proposed approach improves colour neutrality compared with classical statistics-based AWB methods and supervised baselines, while showing promising generalisation from laboratory training conditions to real-world multi-illuminant scenes. These results suggest that incorporating infrared information and semi-supervised learning can improve white balance generalisation and robustness in RGB-IR imaging systems. |
| FP3022-A<br>16:40-16:55 | **Optimality Criteria for Wavelength Allocation in Multi-Camera Multispectral Imaging Systems**<br><br>Author(s): Ivan Škrabo, Ivan Šetka, Josip Vuković, Josip Lončar, Tomislav Matulić, Leon Stjepan Uroić, Jasmin Delihodžić, Dubravko Babić, Damir Seršić,<br>Presenter: Ivan Škrabo, University of Zagreb Faculty of Electrical Engineering and Computing, Croatia<br><br>**Abstract:** While multispectral imaging is typically performed using dedicated multispectral cameras, the high cost of such hardware can be a significant barrier. An alternative involves using standard RGB cameras equipped with narrow-bandpass optical filters. This approach is much cheaper and offers flexibility to swap filters as needed. Such multispectral systems have become well-established in fields such as remote sensing, art conservation and restoration, medical imaging, and food safety inspection [1]–[4]. In this setup, the measured sensor responses are modeled as linear measurements. Each response is determined by the spectral integration of the incoming light, modulated by both the transmittance of the mounted filter and the camera spectral sensitivities. By distributing a predefined set of target wavelengths across multiple camera–filter units, we form a global linear sensing model where the scene spectral irradiance at these wavelengths represents the unknown components to be recovered. The accuracy of this spectral reconstruction depends critically on the numerical stability and properties of the forward operator. Designing filter configurations to achieve stable, noise-resistant imaging has been an active area of research. Hardeberg [5] focused on selecting filters that maximize |

| | |
|---|---|
| | the orthogonality of response vectors, and Li et al. [6] proposed the maximum linear independence criterion, which minimizes the spectral condition number of the sensitivity matrix, thus improving invertibility. These methods treat filter selection as a mathematical problem of how to numerically optimize the sensitivity matrix derived from singular value analysis. Complementary to this perspective, optimal experimental design (OED) theory offers a rich family of design criteria grounded in statistical estimation theory [7]. For a linear forward model with added Gaussian noise, the Fisher Information Matrix (FIM) provides the natural method for quantifying estimation quality. The usual criteria derived from it each offer a unique aspect of reconstruction performance and, generally, lead to different optimal designs. A-optimality targets the average estimation variance by minimizing the trace of the inverse Fisher information matrix, D-optimality maximizes the log-determinant to maximize total information content, and E-optimality promotes worst-case robustness by maximizing the minimum eigenvalue. In this work, we investigate the application of these OED criteria to the problem of wavelength allocation in a multicamera multispectral imaging system. Starting from a physically motivated linear forward model, using a predetermined camera model whose spectral performance is known along with a predefined set of target wavelengths, we construct a global sensing matrix that captures the full interaction between the camera spectral sensitivities and narrow-band filter transmittances. This formulation enables a unified comparison of A-, D-, and E-optimality alongside the condition number and Frobenius norm measures. By systematically evaluating all feasible wavelength allocations under each criterion, we expose the inherent trade-offs between average estimation accuracy, total information gain, and worst-case noise robustness, and identify configurations where different criteria agree or diverge in their recommendations. We demonstrate the framework on a representative setup with twelve target wavelengths and four triband filters mounted on RGB cameras, for which the full space of feasible allocations can be enumerated exhaustively.<br>Table I presents the optimal wavelength sets selected from the predefined candidates for each criterion. This comparison illustrates where the criteria agree and where their priorities differ. Ultimately, this framework provides a sensor-agnostic basis for multispectral system design, offering a flexible and rigorous approach to optimizing diverse camera-filter combinations for specific imaging tasks. |
| FP3024<br>16:55-17:10 | **ROI-Guided Image Coding in JPEG XL**<br><br>Author(s): Hui Wang, Zixi Luo, Claire Mantel, Søren Forchhammer,<br>Presenter: Hui Wang, Technical University of Denmark, Denmark<br><br>**Abstract:** This paper proposes a region-of-interest (ROI)-guided perceptual adaptive quantization framework, enabling spatially adaptive bit allocation without modifying the bitstream syntax or decoder.Within this framework, a perceptual weighting map is derived from a binary ROI mask and integrated into the quantization process to guide bit allocation. Two strategies are investigated for |

| | |
|---|---|
| | perceptual weighting map: per region scaling (M1) and pixel-level weighting (M2). Experimental results show that the proposed framework with both methods, preserve the global rate distortion performance while achieving significant perceptual improvements. Compared to the JPEG XL baseline on a dataset of 30 images, M1 and M2 achieve BD-rate reductions of 6.12% and 7.14% respectively in FVVDP. Subjective testing with 15 participants comparing M1 to the JPEG-XL baseline indicates a strong perceptual preference, with an average success rate of 79.1%. The proposed framework offers a lightweight and practical solution for ROI-based image compression compatible with JPEG XL. |
| FP4042<br>17:10-17:25 | Tiny-Cc-Detector: Localizing Color Charts Using Small Convolutional Neural Networks<br><br>Author(s): Roko Franetović, Marko Subašić<br>Presenter: Roko Franetović, University of Zagreb Faculty of Electrical Engineering and Computing, Croatia<br><br>**Abstract:** Color charts are often used for the task of color calibration, allowing us to accurately calibrate many of our digital cameras. Therefore, color chart localization is an important problem in the domain of colorimetry. In this work we explore small convolutional neural networks for this task with the main goal of reducing the number of model's parameters while maintaining similar performance. In our approach, we transformed the task of color chart localization from regression of bounding box corners into a regression of affine transformation parameters which simplified the original task and allowed us to use smaller models. We developed our synthetic dataset for the modified regression task and trained 6 baseline models which we used for comparison. A grid search of 81 small convolutional models was performed and we compared the top 10 models with our baseline models. We observe that the performance of the models does not scale linearly with the number of parameters and, with respect to the total number of model parameters, we achieve comparable performance with small convolutional models, confirming our hypothesis that our task modification simplified the color chart localization task. |
| FP4041<br>17:25-17:40 | **Variational Inference for a Dynamic Mixture Model**<br><br>Author(s): San Gultekin<br>Presenter: San Gultekin, Meta Inc. USA<br><br>**Abstract:** We introduce a principled way of learning a mixture of predictors for time-varying data using approximate Bayesian inference. Unlike kernel based methods, we derive the model in primal space via Bayesian linear regression. In general, a mixture model from such regressors does not have a closed form posterior, and the mixing parameters are difficult to estimate. We show how variational inference with auxiliary variables can solve the problem, yielding simple updates for |

| | |
|---|---|
| | all parameters involved, which is also guaranteed to find a stationary point of the evidence lower bound. We show results in simulated and real time series. |
| FP506-A<br>17:40-17:55 | **Slope-Track: Multiple Object Tracking on Ski Slopes**<br><br>Author(s): M'Saydez Campbell, Christophe Ducottet, Damien Muselet, Rémi Emonet<br>Presenter: M'Saydez Campbell，Université Jean Monnet Saint-Etienne, France<br><br>**Abstract:** In this paper, we introduce Slope-Track. Slope-Track is a novel multiple object tracking (MOT) dataset designed to reflect the complexities of real ski slope environments. The dataset has over 96,000 frames collected from 10 different ski resorts under various weather and visibility conditions. Slope-Track addresses significant challenges in slope monitoring, including small object sizes, occlusions, fast and irregular motion, and low appearance consistency. It is densely annotated with bounding boxes and object identities, facilitating the evaluation of detection and tracking algorithms. We analyze the dataset's characteristics comparing it to the existing MOT datasets. The results demonstrate that Slope-Track encapsulates a combination of challenges found in other datasets. Additionally, we benchmark a range of existing tracking algorithms and propose a new module that improves motion-based association by dealing with the specific shape of trajectories along ski slopes. Our results demonstrate that incorporating appearance features can have a mixed impact, depending on how they are used within each tracking algorithm. In contrast, motion-based methods and spatial association strategies show more reliable performance. Overall, we provide a challenging benchmark for evaluating and improving multi-object tracking systems in real-world outdoor environments. |

# Online Session

## Cross-Modal Intelligence and Emerging Applications

- **Time:** 10:00-12:25, June 12 [Toulouse Local Time: UTC+2 (CEST)]
- **Zoom link:** https://us02web.zoom.us/j/86103273769
- Zoom ID: 86103273769
- **Password:** 061012
- **Session Chair:** Paulo Batista, University of Évora, Portugal
- **Papers:** Invited Speaker-Paulo Batista, Invited Speaker-Ali H. Wheeb, Invited Speaker-P. Aruna, FP1007, FP4035, FP502, FP711, FP3021

| | |
|---|---|
| Invited Speaker 10:00-10:20 |  <br>**Paulo Batista**<br>**University of Évora, Portugal**<br><br>***Speech Title: Impact of Information on the Enterprise and the Individual***<br><br>**Abstract:** Presented a new interpretation of Theodore Schellenberg's Three Age Theory. They called attention to the fact that the three phases of archival documents are not separate but, on the contrary, integrated. They argued that these three stages can even be looked at in a segmented way, provided the union between them is ensured. Their great innovation relative to Schellenberg's work lay, precisely, in critiquing the division and separation between the three ages of archival documents. Couture and Rousseau thereby brought together all the phases of the lifecycle of records, from production to dissemination, in opposition to the sterile distinction advocated by traditional archivists and document managers. In my opinion, however, the best approach to integrating information management is known as records continuum, which place archives in a post-custodial, informational, and scientific paradigm. This Australian concept arose in the 1990s amid the huge explosion of information, communication technologies and new media. This context forced Information Science to redefine its object of study. Records continuum is closely related to the integrated management model of Couture and Rousseau, while it carries their innovation further, perfecting it and replacing it with systemic dynamics and providing continuity between archives. In fact, records continuum means, literally, continuous management. It looks at the whole process from the production of records to their final archiving. Otherwise, we cannot speak of continuous management. That is why, when we speak of rigid archives – current, intermediate, and definitive, this approach is more theoretical than practical. There is, in fact, no separation between these phases, even less so from the point of view of the value of documents. The traditional distinction between information with probative and historical value ceases to exist. The information is simultaneous and is, in fact, the same. |

| | |
|---|---|
| Invited Speaker 10:20-10:40 |  <br>**Ali H. Wheeb**<br>**University of Baghdad , Baghdad**<br><br>***Speech Title: Intelligence in the Sky: Reinforcement Learning-Driven Routing for Autonomous UAV Networks***<br><br>**Abstract:** The deployment of Unmanned Aerial Vehicle (UAV) networks and Flying Ad-Hoc Networks (FANETs) is critical for time-sensitive, high-mobility applications such as disaster response, environmental monitoring, and autonomous surveillance. However, the operational reality of these networks presents severe challenges: highly dynamic topologies, frequent link breakages, and limited power constraints. Conventionally, static routing protocols are increasingly inadequate for handling such unpredictable flying environments. This presentation introduces a cutting-edge paradigm that leverages Artificial Intelligence (AI) to achieve resilient, autonomous flying communications. Explore the integration of Reinforcement Learning (RL) into routing protocol design, allowing UAV nodes to dynamically learn optimal data paths and adapt to real-time topology changes on the fly. Furthermore, bridge the gap between abstract network algorithms and physical flight constraints by incorporating realistic, aerodynamically bounded mobility models such as 3DGM into the design process. By combining AI-driven adaptability with physical mobility constraints, this approach significantly enhances throughput, reduces latency, and optimizes energy consumption across the swarm. The speech will detail the architecture of these novel RL-routing mechanisms, explain the integration of simulation tool and AI tool to evaluate the performance, and discuss the future of fully autonomous, self-healing aerial networks. |
| Invited Speaker 10:40-11:00 |  <br>**R. Aruna**<br>**Coimbatore Institute of Technology,**<br><br>***Speech Title: Big Data Analytics in Cloud Computing***<br><br>**Abstract:** Big data analytics in the cloud offers scalability, cost-efficiency, speed, and agility. It empowers organizations with advanced analytics capabilities, global accessibility, robust security, and compliance. Embracing cloud-based analytics enables businesses to harness the full potential of their data for strategic decision-making and innovation. |

| | |
|---|---|
| FP1007<br>11:10-11:25 | **Bridging Silence: Visual-Only Hindi Speech Reconstruction using Multimodal AVSR**<br><br>Author(s): Meera Reji, Meeraja K, Meghana Aithal, Megha JS, Shylaja SS<br>Presenter: Meera Reji, PES University, India<br><br>**Abstract:** The recovery of intelligible speech from silent video remains a challenging problem due to the ambiguity of visual speech cues and the difficulty of extracting reliable spatiotemporal information under real-world conditions. This paper presents a context-aware multimodal methodology for lip-to-speech generation in the Hindi language, using strong prior knowledge about the structure of articulation and language-specific phonetic patterns obtained by large-scale pre-training. Unlike previous methods, which rely on conventional visual backbones to perform isolated frame-level feature extractions, our method incorporates a time-aware visual encoder to capture fine-grained lip dynamics and utilizes a pre-trained Hindi audio teacher in order to inject phonetic structure into the process and further stabilize learning. To facilitate effective Hindi sentence reconstruction, predictions at the word level are consolidated at the sentence level, while the effect of visual ambiguities within the video is eliminated using the edit distance refinement technique and prompt tuning of the LLM. Extensive experiments show the superiority of our framework against other existing approaches in terms of transcription fidelity, contextual coherence, and reconstructed Hindi speech in a more natural way. Hence, these results underscore the effectiveness of the approach proposed here and point out its potential applications in assistive communication, silent-speech interfaces, and automated media dubbing. |
| FP4035<br>11:25-11:40 | **Automated Ceramic-Glass Separation Using a Hyperspectral Camera**<br><br>Author(s): Ane Moraza, Fátima Aurora Saiz, Hugo Álvarez<br>Presenter: Ane Moraza. Vicomtech, Spain<br><br>**Abstract:** Ceramics represent one of the most problematic contaminants in glass recycling due to their visual similarity to glass and distinct thermal properties, which lead to product defects and increased costs. This study presents an automated separation system based on Hyperspectral Imaging (HSI). A benchmarking analysis of various preprocessing and machine learning pipelines, using data acquired with a Newtec Buteo hyperspectral camera (430-1700 nm), demonstrated that the combination of Min-Max normalization, Linear Discriminant Analysis (LDA) and Support Vector Machine (SVM) achieves the most accurate classification strategy, reaching a 97.46% success rate. Key spectral features were identified in the 450–570 nm and 1000–1200 nm ranges. The system demonstrates real-time capability, with a processing speed of 0.0806 seconds per line, offering a robust solution for industrial recycling environments. Index Terms—hyperspectral |

| | |
|---|---|
| | imaging, spectral classification, ceramic-glass separation, machine learning, industrial recycling. |
| FP502<br>11:40-11:55 | **Entreacte Inter-Omnia-Glypha Computer Vision & OCR Method**<br><br>Author(s): Luís Homem<br>Presenter: Luís Homem, University of Lisbon, Portugal<br><br>**Abstract:** This paper introduces Entreacte, a novel softwarebased Optical Character Recognition (OCR) and Computer Vision (CV) method founded on the principle of reversed binarization and the generation of differential interval objects termed ”differance ´ ” characters. The method departs from conventional human-centered OCR paradigms by proposing a machine-centered approach that computes the void intervals between characters or visual objects as new computational entities, thereby enhancing both processing efficiency and semantic depth. By integrating morphological operations such as thinning and skeletonization with reversed pixel mapping, Entreacte enables higher-dimensional analysis across text, image, and video domains. Conceptually inspired by Jacques Derrida’s notions of differance ´ and Entreacte, the invention establishes a bridge between philosophical semiotics and computational logic, reinterpreting absence and interstitial space as sources of structural meaning and machine-readable information. This approach expands traditional binarization and segmentation frameworks into a unified paradigm applicable to advanced AI, cryptography, and neural architectures. The resulting system achieves substantial gains (up to 4 × and more) in processing speed, memory compression, and cross-modal recognition, hopefully positioning Entreacte as a fundamental methodological advancement at the intersection of philosophy, information theory, and artificial intelligence. |
| FP711<br>11:55-12:10 | **Analysis on the Development of FitMySpace: A Mobile Application for Furniture Recommendations Using Image-Based Assessment**<br><br>Author(s): Nathan Angelo Mallari, Paulo Rommel Paulino, Kristine Mae Veneracion, Mary Jane Samonte<br>Presenter: Nathan Angelo Mallari, Mapúa University, Philippines<br><br>**Abstract:** This paper presents the analysis of FitMySpace mobile application, designed to enhance furniture shopping through image-based assessment integrated with artificial intelligence (AI) and augmented reality (AR). The study aims to determine user perceptions regarding the application’s usability, efficiency, and overall functionality in improving the online furniture shopping experience. A descriptive quantitative method was employed through an online survey participated in by thirty-one respondents. Findings indicate that integrating AI-driven recommendations and AR visualization can significantly improve user engagement, decision-making, and satisfaction in digital retail. The positive response from participants highlights the potential of FitMySpace to transform the |

| | |
|---|---|
| | furniture industry through intelligent, user-centered design and demonstrates the growing importance of AI-AR integration in modern information technology applications. |
| FP3021<br>12:10-12:25 | **Detection, classification, and alert generation for DDoS attacks in software-defined networks (SDN) using machine learning**<br><br>Author(s): Mamadou Baïlo Bah, Abdoulaye Diallo, Chérif Diallo<br>Presenter: Mamadou Bailo Bah,Université Gaston Berger, Sénégal<br><br>**Abstract:** Software-Defined Networking (SDN) has emerged as a flexible and programmable networking paradigm, but its centralized control architecture exposes critical vulnerabilities to Distributed Denial of Service (DDoS) attacks. Detecting and mitigating such attacks in SDN environments remains a significant challenge due to dynamic traffic patterns and evolving attack strategies.This paper proposes a deployment-oriented intelligent intrusion detection system for DDoS detection, classification, and alert generation in SDN environments using machine learning techniques.The proposed framework integrates an end-to-end pipeline including data preprocessing, correlation-based feature selection, model training, and rigorous evaluation using stratified cross-validation and independent holdout testing. Five supervised learning models are systematically compared under a unified experimental setup.Beyond conventional classification, the framework introduces a probabilistic alert generation mechanism that converts prediction outputs into actionable security alerts based on configurable confidence thresholds. This enables a controllable trade-off between detection performance and false alarm rate, which is essential for real-world operational deployment. Experimental results on a benchmark SDN dataset demonstrate that feature selection improves generalization and enhances minorityclass detection. Among the evaluated models, Random Forest achieves the best overall performance in terms of accuracy, F1-score, and area under the curve. The proposed alerting strategy further improves the reliability of intrusion detection by reducing false positives while maintaining high detection coverage. Overall, the proposed system bridges the gap between theoretical machine learning-based detection and practical intrusion detection systems by integrating detection, alerting, and response capabilities into a unified, scalable, and reproducible framework suitable for real-world SDN environments. |

# Notes