Xavier Alameda-Pineda, Eng., Ph.D., HDR

Research Director and Leader of the RobotLearn
Inria @ Univ. Grenoble Alpes, MIAI

xavier.alameda-pineda@inria.fr
Scholar LinkedIn Bluesky

(last updated on June 2026)

Academic and Professional Qualifications

2024 – **ELLIS Fellow** (ELLIS, Europe)
2020 – **Habilitation à Diriger de Recherches [TH1]** (UGA, Grenoble, France)
2019 – **IEEE Senior Member** (IEEE, USA)
2013 – **PhD in Mathematics and Computer Science [TH2]** (Perception Team, Inria, UJF, Grenoble, France)
2010 – **Masters in Computer Science (Graphics, Vision and Robotics) [TH3]** (MoSIG, UFR IM2AG, UJF, Grenoble-INP, Grenoble, France)
2009 – **Masters (equiv.) in Telecommunications Engineering [TH4]** (ETSETB, BarcelonaTECH, Barcelona, Spain)
2008 – **Masters (equiv.) in Mathematics** (FME, BarcelonaTECH, Barcelona, Spain)

Research Experience

2024–Now – **Research Director** (RobotLearn Team, Inria, Grenoble, France)
2021–Now – **Research Team Leader** (RobotLearn Team, Inria, Grenoble, France)
2016–2024 – **Research Scientist** (RobotLearn (prev. Perception) Team, Inria, Grenoble, France)
2014–2016 – **Postdoctoral Fellow** [with Prof. Nicu Sebe] (MHUG, DISI, University of Trento, Trento, Italy)
2013–2014 – **Postdoctoral Fellow** [with Profs. Laurent Girin & Radu Horaud] (MAGIC/Perception Teams, GIPSA-LAB/Inria, Grenoble, France)
2010–2013 – **PhD Candidate (French State Funding)** [with Dr. Radu Horaud] (Perception Team, Inria, UJF, Grenoble, France)
2007–2008 – **Research collaboration grant** [with Prof. Philippe Salembier] (IPG, STC, UPC, Barcelona, Spain)
2007–2008 – **Research collaboration grant for young researchers** [with Prof. Xavier Cabré] (MEC and MA1, UPC, Barcelona, Spain)

Awards and Honors

IEEE TMM Outstanding Associate Editor Award for *Contributions to the IEEE TMM Editorial Board*
at IEEE Transactions on Multimedia
ACM TOMM Nicolas D. Georganas Best Paper Award for *"Increasing image memorability with neural style transfer"* [J30]
at ACM Transactions on Multimedia Computing
ACM SIGMM Rising Star Award for *Contributions to Multimodal Social Behavior Understanding*
at ACM SIGMM
IEEE IROS Novel Technology Award Finalist for *Tracking a Varying Number of People with a Visually-Controlled Robotic Head* [C41]
at IEEE IROS 2017
IAPR ICPR Best Scientific Paper Award for *Multi-Paced Dictionary Learning for Cross-Domain Retrieval and Recognition* [C52]
at IAPR ICPR 2016
ACM MM Best Paper Award for *Analysing Free-standing Conversational Groups: A Multimodal Approach* [C53]
at ACM International Conference on Multimedia 2015
IEEE WASPAA Best Student Paper Award for *A Variational EM Algorithm for the Separation of Moving Sound Sources* [C54]
at IEEE WASPAA 2015
ACM ICMI Outstanding Paper Award for *Finding Audio-visual Events in Informal Social Gatherings* [C64]
at IEEE/ACM ICMI 2011

Funding

H2020 – Coordinator [2020–2024, 8.3 M€] SPRING: Socially pertinent robots for gerontological healthcare
ANR-JCJC & IDEX-IRS – PI [2020–2024, 310 k€] ML3RI: Multi-modal multi-person low-level learning for robot interaction
ANR 3IA – co-PI [2019–2023, 360 k€] MIAI Chair: Audio-visual machine perception and interaction for companion robots
IDEX-ISP – PI [2019–2023, 45 k€] PIMPE: Physical complex Interactions and Multi-person Pose Estimation
IDEX-IRS – co-PI [2019–2023, 90 k€] MIDGen: Multimodal Interaction Data Generation

Supervision

Name	Period (Affiliation)	Status & Publications	Co-Supervisor
PhD (co-)Supervision – (co-)Directeur de thèse			
Maxime Attwood	2025– (Inria Grenoble)	Ongoing	Dr. Stéphane Lathuilière
Sofiene Kammoun	2024– (Central Supélec)	Ongoing [C2]	Prof. Simon Leglaive
Jordan Cosio	2023– (Inria Grenoble)	Ongoing [C9]	Dr. Pierre-Brice Wieber
Jean-Eudes Ayilo	2023– (Inria Nancy)	Ongoing [C4, J1, J5]	Dr. Mostafa Sadeghi & Prof. Romain Serizel
Gaétan Lepage	2020–25 (Inria Grenoble)	Defended (17/07/2025)	Prof. Laurent Girin & Dr. Chris Reinke
Guénolé Fiche	2023–24 (Central Supélec)	Defended (18/11/2024) [C11, C7, C5]	Prof. Simon Leglaive
Samir Sadok	2023–24 (Central Supélec)	Defended (08/03/2024) [J14, J8]	Prof. Simon Leglaive
Xiaoyu Lin	2020–24 (Inria Grenoble)	Defended (25/06/2024) [C14, J12, C15]	Prof. Laurent Girin
Anand Ballou	2019–24 (Inria Grenoble)	Defended (27/03/2024) [J10]	Dr. Chris Reinke
Louis Airale	2019–23 (Inria Grenoble)	Defended (04/12/2023) [J15, J6]	Prof. Dominique Vaufreydaz
Xiaoyu Bie	2019–23 (Inria Grenoble)	Defended (20/10/2023) [C23, J17, J16, C19]	Prof. Laurent Girin
Wen Guo	2019–2023 (Inria Grenoble)	Defended (12/06/2023) [C24, C19, C13]	Dr. Francesc Moreno-Noguer
Guanglei Yang	2021–22 (HIT)	Defended (12/06/2023) [J19, J20]	Prof. Elisa Ricci
Enrico Fini	2021–22 (U. Trento)	Defended (17/04/2023) [C17, C12]	Prof. Elisa Ricci & Dr. Karteek Alahari
Hanyu Xuan	2021–22 (NUST)	Defended (2022) [C21, C22]	Prof. Yan Yan
Yihong Xu	2018–22 (Inria Grenoble)	Defended (08/06/2022) [C32, J18, C22, C28]	Dr. Radu Horaud
Guillaume Delorme	2017–21 (Inria Grenoble)	Defended (08/10/2021) [C28, W3, J18, C33]	Dr. Radu Horaud
Aliaksandr Siarohin	2016–17 (U. Trento)	Defended (24/06/2021) [C45, J30]	Prof. Nicu Sebe
Andrea Pilzer	2016–17 (U. Trento)	Defended (22/06/2020) [C39]	Prof. Nicu Sebe
Yutong Ban	2016–19 (Inria Grenoble)	Defended (10/05/2019) [W6, W7, C41, J27, J22, C36]	Dr. Radu Horaud
Wei Wang	2016–18 (U. Trento)	Defended (2018) [C50, C38, J25]	Prof. Nicu Sebe
Stéphane Lathuilière	2016–17 (Inria Grenoble)	Defended (22/05/2018) [J28, C37, C28]	Dr. Radu Horaud
Israel-Dejene Gebru	2014–17 (Inria Grenoble)	Defended (13/04/2018) [C55, J36]	Dr. Radu Horaud
Dan Xu	2014–17 (U. Trento)	Defended (2018) [C46, J26, J31, C51, C52]	Prof. Nicu Sebe
Sergey Tulyakov	2015–17 (U. Trento)	Defended (2017) [C49]	Prof. Nicu Sebe
Vincent Drouard	2016–17 (Inria Grenoble)	Defended (18/12/2017) [J21]	Dr. Radu Horaud
Dionyssos Kounades-Bastian	2014–17 (Inria Grenoble)	Defended (24/02/2017) [J37, C48, C43, C54, C44]	Dr. Radu Horaud & Prof. Laurent Girin
MSc (co-)Supervision – (co-)Encadrant de Stage			
Maxime Attwood	2025 (Inria Grenoble)	Defended (2025)	Dr. Samir Sadok
Ahmad Ghazi	2024 (Inria Grenoble)	Defended (2024)	Dr. Pedro Rodrigues
David Emukpere	2021 (Inria Grenoble)	Defended (2021)	Dr. Chris Reinke
Viet Nhat Nguyen	2020 (Inria Grenoble)	Defended (2020)	Dr. Mostafa Sadeghi
Álvaro González Jiménez	2020 (Inria Grenoble)	Defended (2020)	Prof. Stéphane Lathuilière
Predrag Pilipovic	2020 (Inria Grenoble)	Defended (2020)	
Vadim Sushko	2019 (Inria Grenoble)	Defended (2019)	
Elsa Marie	2019 (Inria Grenoble)	Defended (2019)	Prof. Simon Leglaive
Divya Grover	2018 (Inria Grenoble)	Defended (2018)	Dr. Radu Horaud
Alessio Xompero	2014 (Inria Grenoble)	Defended (2014)	Dr. Radu Horaud
Israel-Dejene Gebru	2013 (Inria Grenoble)	Defended (2013)	Dr. Radu Horaud
Maxime Janvier	2012 (Inria Grenoble)	Defended (2012)	Dr. Radu Horaud

Teaching

2025–2026 – **Generative Multimodal AI (5h)** [with Dr. Stéphane Lathuilière, Dr. Pia Bideau, Dr. Eric Gaussier] (ENSIMAG/UGA, Grenoble, France)
2025–2026 – **Learning, Probabilities and Causality (5h)** [with Dr. Thomas Hueber, Dr. Lorena León, Dr. Eric Gaussier] (ENSIMAG/UGA, Grenoble, France)
2024–2025 – **Advanced Machine Learning: Applications to Vision, Audio and Text (12h)** [with Dr. Karteek Alahari, Dr. Eric Gaussier] (ENSIMAG/UGA, Grenoble, France)
2023–2024 – **Learning, Probabilities and Causality (15h)** [with Dr. Thomas Hueber, Dr. Eric Gaussier] (ENSIMAG/UGA, Grenoble, France)
2023–2024 – **Advanced Machine Learning: Applications to Vision, Audio and Text (12h)** [with Dr. Karteek Alahari, Dr. Eric Gaussier] (ENSIMAG/UGA, Grenoble, France)
2022–2023 – **Learning, Probabilities and Causality (15h)** [with Dr. Thomas Hueber, Dr. Eric Gaussier] (ENSIMAG/UGA, Grenoble, France)
2022–2023 – **Advanced Machine Learning: Applications to Vision, Audio and Text (12h)** [with Dr. Karteek Alahari, Dr. Eric Gaussier] (ENSIMAG/UGA,

Grenoble, France)

2021–2022 — **Fundamentals of Probabilistic Data Mining (15h)** [with Dr. Thomas Hueber] (ENSIMAG/UGA, Grenoble, France)
2021–2022 — **Machine Learning for Multimodal Data (12h)** [with Dr. Karteek Alahari, Dr. Eric Gaussier] (ENSIMAG/UGA, Grenoble, France)
2020–2021 — **Fundamentals of Probabilistic Data Mining (15h)** [with Dr. Thomas Hueber] (ENSIMAG/UGA, Grenoble, France)
2020–2021 — **Machine Learning for Computer Vision and Audio Processing (12h)** [with Dr. Karteek Alahari] (ENSIMAG/UGA, Grenoble, France)
2019–2020 — **Category Learning and Object Recognition (9h)** [with Dr. Karteek Alahari] (ENSIMAG/UGA, Grenoble, France)
2019–2020 — **Advanced Learning Models (9h)** [with Dr. Julien Mairal] (ENSIMAG/UGA, Grenoble, France)
2019–2020 — **Fundamentals of Probabilistic Data Mining (13h)** [with Dr. Thomas Hueber] (ENSIMAG/UGA, Grenoble, France)
2018–2019 — **Fundamentals of Probabilistic Data Mining (5h)** [with Prof. Jean-Baptiste Durand] (ENSIMAG/UGA, Grenoble, France)
2012–2013 — **Algebra and Geometry (128h)** [with Profs. François Dahmani and Stéphanie Pouget] (DLST, UJF, Grenoble, France)

Community Service

Technical Steering Committees: ACM SIGMM Multimedia Steering Committee (2022–); IEEE TC on Audio Acoustics and Signal Processing (2022–2024); IAPR TC9 on Pattern Recognition in Human Machine Interaction (2021–2023).

Award Committees: Member: ACM SIGMM Rising Star Award 2025, Member: ACM Multimedia Best Paper Award 2025, co-Chair: IEEE ICME Best Paper Award Committee 2024, co-Chair: IEEE TMM Best Paper Award Committee 2023.

Associated Editor (period, # handled papers): TMRL (2026–, 2), CVIU (2021–, 14), ACM TOMM (2019–2025, 41), IEEE TMM (2021–2024, 55), ACM TIST (2021–2024, 11).

Conference Leadership: General Chair: ACM MM 2026; Program Chair: ACM MM 2022; Senior Area Chair: ACM MM 2025 (215 papers); Senior Area Chair: ACM MM 2024 (185 papers).

Area Chair (# handled papers): IEEE ICASSP 2025 (20), ICIAP 2025 (8), ICRA 2024 (11), ACM MM 2023 (25), IEEE WASPAA 2023 (4), IEEE ICASSP 2023 (12), ICRA 2023 (11), BMVC 2022 (8), AAAI 2022 (2), ACM MM 2021 (9), IEEE WACV 2021 (20), ACM MM 2020 (26), IAPR ICPR 2020 (20), ACM MM 2019 (24), ICIAP 2019 (17), ICCV 2017 (29).

Workshop Organiser: T-CAP @ ECCV 2024, Meet4MM @ ACM MM 2024, WCPA @ ECCV 2022, M4MM @ ACM MM 2022, MPRSS @ IAPR ICPR 2022, FATE-MM @ ACM MM 2020, MediaEval @ 2019, WHBU @ ICCV 2019, FAT-MM @ ACM MM 2019, FFSS-USAD @ CVPR 2019, HBUGEN @ ECCV 2018, EE-USAD @ ACM MM 2018, V-USAD @ CVPR 2018.

Tutorials & Invited Talks

Multimodal perception, action, and evaluation of socially intelligent robots – invited (*Nov'25*)
at International Workshop on AI for Robotics, Meylan, FR
Audio-visual speech processing with probabilistic models – invited (*Oct'25*)
at MILA, Montréal, CA
From VAE to Diffusion – probabilistic learning with audio-visual data – course (*Sep'25*)
at INPT AI Summer School, Rabat, MA
Probabilistic generative models for audio-visual processing – course (*Oct'24*)
at Paris GenAI Autumn School, Saclay, FR
Learning for Companion Robots: Preparation and Adaptation – keynote – Slides (*Jul'24*)
at Joint RFIAP-cAP, Lille, FR
Social Robot Learning – invited (*Jun'24*)
at CEA List Days, Saclay, FR
Variational Audio-Visual Representation Learning – keynote – Slides (*Nov'23*)
at ACM International Conference on Multimedia, Ottawa, CA
Learning for Robots in Conversational Groups – invited (*May'23*)
at Workshop of the International Laboratory on Learning Systems, Université Paris Saclay
Robots within Groups of People – invited (*May'23*)
at Interdisciplinary Workshop on Mingling Technologies, TU Delft
Unsupervised Probabilistic Learning with Latent Variables – course – Slides (*Jan'23*)
at Machine Learning Summer School Africa 2023
Learning for Socially Intelligent Robots – invited (*Dec'22*)
at Computer Science and Electric Engineering Departments, University of Alberta
Introduction to Dynamical Variational Autoencoders – invited (*Feb'22*)
at MLIA Research Team Seminars
Deep Generative Modeling of Sequential Data with Dynamical Variational Autoencoders – tutorial – Recording (*Jun'21*)
at IEEE ICASSP
Unsupervised Learning for Human Robot Perception – invited – Recording (*Jun'21*)

at Robotics and AI Summer School 2021

Towards socially intelligent robots: preliminary results of the H2020 SPRING and the ANR ML3RI projects – invited – Recording (*Jun'21*)
at PI Stories University of Trento

Unsupervised Audio-Visual Fusion for Upstream Human Behavior Understanding – invited – Recording (*May'21*)
at AI4Media Workshop on New Learning Paradigms and Distributed AI

Variational Autoencoders for Audio, Visual and Audio-Visual Learning – tutorial – Recording (*Feb'21*)
at DaSCI Webinars

Speaker localisation and enhancement in populated environments – invited (*Jan'21*)
at ICPR 2020 Workshop on Deep Learning for Human-Centric Activity Understanding

Combining auditory and visual data to enhance the speech signal – invited (*Jan'21*)
at ICPR 2020 Workshop on Multimodal Pattern Recognition for Social Signal Processing

Towards audio-visual speech enhancement in robotic platforms – invited (*Dec'20*)
at Journée GT5 Interactions Personnes/Systèmes Robotiques, GDR Robotique

Audio-visual variational speech enhancement – tutorial (*Sep'20*)
at Intelligent Sensing Summer School

Choosing wisely your deep training loss – invited (*Mar'20*)
at Universidade NOVA de Lisboa

Artificial Intelligence for Social Robots in Gerontological Healthcare – invited (*Mar'20*)
at European Robotics Forum

Probabilistic and deep learning for regression in computer vision – tutorial (*Sep'19*)
at ICIAP 2019

Significancy & Robustness in Deep Regression – invited (*Jul'19*)
at University of Trento

Probabilistic and deep methods for human behavior understanding – invited (*Jul'19*)
at Media Integration and Communication Center

Multi-speaker audio-visual diarization – invited (*Dec'18*)
at SOUND Workshop Bar-Ilan

Multimodal social behavior understanding – invited (*Oct'18*)
at ACM SIGMM Rising Star Lecture at ACM MM

Audio-Visual Multiple Speaker with Robotic Platforms – invited (*May'18*)
at University of Trento and RHUM Workshop

Multimodal human behavior analysis in the wild – tutorial (*Dec'16*)
at IAPR ICPR 2016

Emerging topics in noisy and missing data – tutorial (*Oct'16*)
at ACM MM 2016

Matrix completion: a computer vision perspective – invited (*Jun'16*)
at Carnegie Mellon University & Digital Video and Multimedia Lab, Columbia University

Multimodal behavioral signal processing in the wild – invited (*Jun'16*)
at Télécom-ParisTech

Variational EM and non-linear optimization for multi-sensor scene analysis – invited (*Dec'15*)
at Laboratoire d'Analyse et d'Architecture de Systèmes, CNRS

Free-standing conversational groups: the SALSA dataset and multi-modal head and body pose estimation – invited (*Nov'15*)
at UPC Image and Video Processing Group & Inria Nancy Team multispeech

Multimodal Automatic Analysis of Group Behavior – invited (*Oct'15*)
at SIGMM Inaugural Workshop on Multimedia Frontiers

Publications

Journal Articles

[J1] Jean-Eudes Ayilo, Mostafa Sadeghi, Romain Serizel, and Xavier Alameda-Pineda. "Diffusion-based Frameworks for Unsupervised Speech Enhancement". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2026).

[J2] Xavier Alameda-Pineda et al. "Socially Pertinent Robots in Gerontological Healthcare". In: *International Journal on Social Robotics* (2025). Under Review at International Journal on Social Robotics. DOI: https://doi.org/10.1007/s12369-025-01330-6.

[J3] Lauriane Blavette, Sébastien Dacunha, Xavier Alameda-Pineda, Jeanne Cattoni, Anne-Sophie Rigaud, and Maribel Pino. "Integrating a Large Language Model Into a Socially Assistive Robot in a Hospital Geriatric Unit: Two-Wave Comparative Study on Performance, Engagement, and User Perceptions". In: *JMIR Human Factors* 12.1 (2025), e81936.

[J4] Lauriane Blavette, Sébastien Dacunha, Xavier Alameda-Pineda, Daniel Hernández García, Sharon Gannot, Florian Gras, Nancie Gunson, Séverin Lemaignan, Michal Polic, Pinchas Tandeitnik, et al. "Acceptability and usability of a socially assistive robot integrated with a large language model for enhanced human-robot interaction in a geriatric care institution: mixed methods evaluation". In: *JMIR Human Factors* (2025).

[J5] Mostafa Sadeghi, Jean-Eudes Ayilo, Romain Serizel, and Xavier Alameda-Pineda. “Posterior Transition Modeling for Unsupervised Diffusion-Based Speech Enhancement”. In: *IEEE Signal Processing Letters* (2025).

[J6] Louis Airale, Xavier Alameda-Pineda, Stéphane Lathuilière, and Dominique Vaufreydaz. “Autoregressive GAN for Semantic Unconditional Head Motion Generation”. In: *ACM Transactions on Multimedia Computing, Communications, and Applications* (2024). DOI: 10.1145/3635154.

[J7] Mostafa Sadeghi, Xavier Alameda-Pineda, and Radu Horaud. “Unsupervised performance analysis of 3D face alignment with a statistically robust confidence test”. In: *Neurocomputing* 564 (2024). https://team.inria.fr/robotlearn/upa3dfa/. DOI: 10.1016/j.neucom.2023.126941.

[J8] Samir Sadok, Simon Leglaive, Laurent Girin, Xavier Alameda-Pineda, and Renaud Séguier. “A Multimodal Dynamical Variational Autoencoder for Audiovisual Speech Representation Learning”. In: *Neural Networks* (2024). https://samsad35.github.io/site-mdvae/.

[J9] Hanyu Xuan, Zhiliang Wu, Jian Yang, Bo Jiang, Lei Luo, Xavier Alameda-Pineda, and Yan Yan. “Robust Audio-Visual Contrastive Learning for Proposal-based Self-supervised Sound Source Localization in Videos”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46 (2024), pp. 4896–4907. DOI: 10.1109/TPAMI.2024.3363508.

[J10] Anand Ballou, Xavier Alameda-Pineda, and Chris Reinke. “Variational Meta Reinforcement Learning for Social Robotics”. In: *Applied Intelligence* 53 (2023), pp. 27249–27268. DOI: 10.1007/s10489-023-04691-5.

[J11] Zhiqi Kang, Mostafa Sadeghi, Radu Horaud, and Xavier Alameda-Pineda. “Expression-preserving face frontalization improves visually assisted speech processing”. In: *International Journal of Computer Vision* 131 (5 2023), pp. 1122–1140. DOI: 10.1007/s11263-022-01742-1.

[J12] Xiaoyu Lin, Laurent Girin, and Xavier Alameda-Pineda. “Mixture of Dynamical Variational Autoencoders for Multi-Source Trajectory Modeling and Separation”. In: *Transactions on Machine Learning Research* (2023).

[J13] Chris Reinke and Xavier Alameda-Pineda. “Successor Feature Representations”. In: *Transactions on Machine Learning Research* (2023).

[J14] Samir Sadok, Simon Leglaive, Laurent Girin, Xavier Alameda-Pineda, and Renaud Séguier. “Learning and controlling the source-filter representation of speech with a variational autoencoder”. In: *Speech Communication* 148 (Mar. 2023). https://samsad35.github.io/site-sfvae/, pp. 53–65. DOI: 10.1016/j.specom.2023.02.005.

[J15] Louis Airale, Dominique Vaufreydaz, and Xavier Alameda-Pineda. “SocialInteractionGAN: Multi-person Interaction Sequence Generation”. In: *IEEE/ACM Transactions on Affective Computing* (2022). DOI: 10.1109/TAFFC.2022.3171719.

[J16] Xiaoyu Bie, Simon Leglaive, Xavier Alameda-Pineda, and Laurent Girin. “Unsupervised Speech Enhancement using Dynamical Variational Auto-Encoders”. In: *IEEE/ACM Transactions on Audio, Signal and Language Processing* (2022). DOI: 10.1109/TASLP.2022.3207349.

[J17] Laurent Girin, Simon Leglaive, Xiaoyu Bie, Julien Diard, Thomas Hueber, and Xavier Alameda-Pineda. “Dynamical Variational Autoencoders: A Comprehensive Review”. In: *Foundations and Trends in Machine Learning* 1-2.15 (2022). DOI: 10.1561/2200000089.

[J18] Yihong Xu, Yutong Ban, Guillaume Delorme, Chuang Gan, Daniela Rus, and Xavier Alameda-Pineda. “TransCenter: Transformers with Dense Queries for Multiple-Object Tracking”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022). DOI: 10.1109/TPAMI.2022.3225078.

[J19] Guanglei Yang, Enrico Fini, Dan Xu, Paolo Rota, Mingli Ding, Tang Hao, Xavier Alameda-Pineda, and Elisa Ricci. “Continual Attentive Fusion for Incremental Learning in Semantic Segmentation”. In: *IEEE Transactions on Multimedia* 25 (2022), pp. 3841–3854. DOI: 10.1109/TMM.2022.3167555.

[J20] Guanglei Yang, Enrico Fini, Dan Xu, Paolo Rota, Mingli Ding, Moin Nabi, Xavier Alameda-Pineda, and Elisa Ricci. “Uncertainty-aware Contrastive Distillation for Incremental Semantic Segmentation”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2022). DOI: 10.1109/TPAMI.2022.3163806.

[J21] Xavier Alameda-Pineda, Vincent Drouard, and Radu Horaud. “Variational Inference and Learning of Piecewise-linear Dynamical Systems”. In: *IEEE Transactions on Neural Networks and Learning Systems* (2021). DOI: 10.1109/TNNLS.2021.3054407.

[J22] Yutong Ban, Xavier Alameda-Pineda, Laurent Girin, and Radu Horaud. “Variational Bayesian Inference for Audio-Visual Tracking of Multiple Speakers”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.5 (2020), pp. 1761–1776. DOI: 10.1109/TPAMI.2019.2953020.

[J23] Mostafa Sadeghi and Xavier Alameda-Pineda. “Mixture of Inference Networks for VAE-based Audio-visual Speech Enhancement”. In: *IEEE Transactions on Signal Processing* 69 (2020), pp. 1899–1909. DOI: 10.1109/TSP.2021.3066038.

[J24] Mostafa Sadeghi, Simon Leglaive, Xavier Alameda-Pineda, Laurent Girin, and Radu Horaud. “Audio-visual Speech Enhancement Using Conditional Variational Auto-Encoders”. In: *IEEE Transactions on Audio, Language and Signal Processing* (2020). DOI: 10.1109/TASLP.2020.3000593.

[J25] Wei Wang, Xavier Alameda-Pineda, Dan Xu, Elisa Ricci, and Nicu Sebe. “Learning How to Smile: Expression Video Generation with Conditional Adversarial Recurrent Nets”. In: *IEEE Transactions on Multimedia* 22.11 (2020), pp. 2808–2819. DOI: 10.1109/TMM.2019.2963621.

[J26] Dan Xu, Xavier Alameda-Pineda, Wanly Ouyang, Elisa Ricci, Xiaogang Wang, and Nicu Sebe. “Probabilistic Graph Attention Network with Conditional Kernels for Pixel-Wise Prediction”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020). DOI: 10.1109/TPAMI.2020.3043781.

[J27] Y. Ban, X. Alameda-Pineda, C. Evers, and R. Horaud. “Tracking Multiple Audio Sources with the Von Mises Distribution and Variational EM”. In: *IEEE Signal Processing Letters* 26.6 (2019), pp. 798–802. DOI: 10.1109/LSP.2019.2908376.

[J28] Stèphane Lathuilière, Pablo Mesejo, Xavier Alameda-Pineda, and Radu Horaud. “A Comprehensive Analysis of Deep Regression”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019). DOI: 10.1109/TPAMI.2019.2910523.

[J29] Xiaofei Li, Yutong Ban, Laurent Girin, Xavier Alameda-Pineda, and Radu Horaud. “Online Localization and Tracking of Multiple Moving Speakers in Reverberant Environment”. In: *IEEE Journal of Selected Topics in Signal Processing* (2019), pp. 88–103. DOI: 10.1109/JSTSP.2019.2903472.

[J30] A Siarohin, G. Zen, C Majtanovic, X. Alameda-Pineda, E. Ricci, and N. Sebe. “Increasing Image Memorability with Neural Style Transfer”. In: *ACM Transactions on Multimedia Computing Communications and Applications* (2019). DOI: 10.1145/3311781.

[J31] Dan Xu, Xavier Alameda-Pineda, Jingkuan Song, Elisa Ricci, and Nicu Sebe. “Cross-Paced Representation Learning with Partial Curricula for Sketch-based Image Retrieval”. In: *IEEE Transactions on Image Processing* (2018). DOI: 10.1109/TIP.2018.2837381.

[J32] Diandra Fabre, Thomas Hueber, Laurent Girin, Xavier Alameda-Pineda, and Pierre Badin. “Automatic Animation of an Articulatory Tongue Model from Ultrasound Images of the Vocal Tract”. In: *Speech Communications* 93 (2017), pp. 63–75. DOI: 10.1016/j.specom.2017.08.002.

[J33] L. Girin, T. Hueber, and X. Alameda-Pineda. “Extending the Cascaded Gaussian Mixture Regression Framework for Cross-Speaker Acoustic-Articulatory Mapping”. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* (2017). DOI: 10.1109/TASLP.2017.2651398.

[J34] Xavier Alameda-Pineda, Jacopo Staiano, Ramanathan Subramanian, Ligia Maria Batrinca, Elisa Ricci, Bruno Lepri, Oswald Lanz, and Nicu Sebe. “SALSA: A Novel Dataset for Multimodal Group Behavior Analysis”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38.8 (2016), pp. 1707–1720. DOI: 10.1109/TPAMI.2015.2496269.

[J35] Sileye Ba, Xavier Alameda-Pineda, Alessio Xompero, and Radu Horaud. “An On-line Variational Bayesian Model for Multi-Person Tracking from Cluttered Scenes”. In: *Computer Vision and Image Understanding* 153 (2016), pp. 64–76. DOI: 10.1016/j.cviu.2016.07.006.

[J36] Israel-Dejene Gebru, Xavier Alameda-Pineda, Florence Forbes, and Radu Horaud. “EM algorithms for weighted-data clustering with application to audio-visual scene analysis”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 38.12 (2016), pp. 2402–2415. DOI: 10.1109/TPAMI.2016.2522425.

[J37] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Sharon Gannot, and Radu Horaud. “A Variational EM Algorithm for the Separation of Time-Varying Convolutive Audio Mixtures”. In: *IEEE/ACM Transactions on Audio, Speech and Language Processing* 24.8 (2016), pp. 1408–1423. DOI: 10.1109/TASLP.2016.2554286.

[J38] Xavier Alameda-Pineda and Radu Horaud. “Vision-Guided Robot Hearing”. In: *International Journal of Robotics Research* 34.4-5 (2015), pp. 437–456. DOI: 10.1177/0278364914548050.

[J39] Thomas Hueber, Laurent Girin, Xavier Alameda-Pineda, and Gerard Bailly. “Speaker-Adaptive Acoustic-Articulatory Inversion using Cascaded Gaussian Mixture Regression”. In: *IEEE/ACM Transactions on Audio, Speech and Language Processing* 23.12 (2015), pp. 2246–2259. DOI: 10.1109/TASLP.2015.2464702.

[J40] Xavier Alameda-Pineda and Radu Horaud. “A Geometric Approach to Sound Source Localization from Time-Delay Estimates”. In: *IEEE Transactions on Audio, Speech and Language Processing* 22.6 (2014), pp. 1082–1095. DOI: 10.1109/TASLP.2014.2317989.

[J41] Xavier Alameda-Pineda, Jordi Sanchez-Riera, Johannes Wienke, Vojtech Franc, Jan Cech, Kaustubh Kulkarni, Antoine Deleforge, and Radu Horaud. “RAVEL: An Annotated Corpus for Training Robots with Audiovisual Abilities”. In: *Journal on Multimodal User Interfaces* 7.1-2 (2013), pp. 79–91. DOI: 10.1007/s12193-012-0111-y.

Conference Papers

[C1] Jordan Cosio and Dhimiter Pikuli, Thibaut Lopez, Xavier Alameda-Pineda, and Pierre-Brice Wieber. “LiSSEN: Generating Language-Grounded Safe Social Navigation in Crowded Environments”. In: *IEEE International Conference on Robot and Human Interactive Communication*. 2026.

[C2] Sofiene Kammoun, Xavier Alameda-Pineda, and Simon Leglaive. “Modeling strategies for speech enhancement in the latent space of a neural audio codec”. In: *ICASSP 2026-2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2026.

[C3] Samir Sadok, Stéphane Lathuilière, and Xavier Alameda-Pineda. “Residual Tokens Enhance Masked Autoencoders for Speech Modeling”. In: *Interspeech*. 2026.

[C4] Jean Eudes Ayilo, Mostafa Sadeghi, Romain Serizel, and Xavier Alameda-Pineda. “Diffusion-based Unsupervised Audio-visual Speech Enhancement”. In: *IEEE International Conference on Acoustics, Speech and Audio Processing*. 2025.

[C5] Guénolé Fiche, Simon Leglaive, Xavier Alameda-Pineda, and Francesc Moreno-Noguer. “MEGA: Masked Generative Autoencoder for Human Mesh Recovery”. In: *IEEE International Conference on Computer Vision and Pattern Recognition*. 2025.

[C6] Samir Sadok, Simon Leglaive, Laurent Girin, Gaël Richard, and Xavier Alameda-Pineda. “AnCoGen: Analysis, Control and Generation of Speech with a Masked Autoencoder”. In: *IEEE International Conference on Audio, Speech, and Signal Processing*. 2025.

[C7] Guénolé Fiche, Simon Leglaive, Xavier Alameda-Pineda, Antonio Agudo, and Francesc Moreno-Noguer. “VQ-HPS: Human Pose and Shape Estimation in a Vector-Quantized Latent Space”. In: *European Conference on Comptuer Vision*. 2024. DOI: 10.1007/978-3-031-72943-0_27.

[C8] Ali Golmakani, Mostafa Sadeghi, Xavier Alameda-Pineda, and Romain Serizel. “A weighted-variance variational autoencoder model for speech enhancement”. In: *IEEE International Conference on Acoustics Speech and Signal Processing*. 2024. DOI: 10.1109/ICASSP48485.2024.10446294.

[C9] Dhimiter Pikuli, Jordan Cosio, Xavier Alameda-Pineda, Pierre-Brice Wieber, and Thierry Fraichard. “Navigating the Practical Pitfalls of Reinforcement Learning for Social Robot Navigation”. In: *Robotics: Science and Systems (RSS) Workshop on Unsolved Problems in Social Robot Navigation*. 2024.

[C10] Lorenzo Vaquero, Yihong Xu, Xavier Alameda-Pineda, Victor M Brea, and Manuel Mucientes. “Lost and Found: Overcoming Detector Failures in Online Multi-Object Tracking”. In: *European Conference on Computer Vision*. 2024. DOI: 10.1007/978-3-031-73464-9_27.

[C11] Guénolé Fiche, Simon Leglaive, Xavier Alameda-Pineda, and Renaud Séguier. “Motion-DVAE: Unsupervised learning for fast human motion denoising”. In: *ACM SIGGRAPH Conference on Motion, Interaction and Games*. 2023. DOI: 10.1145/3623264.3624454.

[C12] Enrico Fini, Pietro Astolfi, Karteek Alahari, Xavier Alameda-Pineda, Julien Mairal, Moin Nabi, and Elisa Ricci. “Semi-supervised learning made simple with self-supervised clustering”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, pp. 3187–3197. DOI: 10.1109/CVPR52729.2023.00311.

[C13] Wen Guo, Yuming Du, Xi Shen, Vincent Lepetit, Xavier Alameda-Pineda, and Francesc Moreno-Noguer. “Back to MLP: A Simple Baseline for Human Motion Prediction”. In: *IEEE Winter Conference on Applications of Computer Vision*. 2023, pp. 4809–4819. DOI: 10.1109/WACV56688.2023.00479.

[C14] Xiaoyu Lin, Simon Leglaive, Laurent Girin, and Xavier Alameda-Pineda. “Speech Modeling with a Hierarchical Transformer Dynamical VAE”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. 2023. DOI: 10.1109/ICASSP49357.2023.10096751.

[C15] Xiaoyu Lin, Simon Leglaive, Laurent Girin, and Xavier Alameda-Pineda. “Unsupervised speech enhancement with deep dynamical generative speech and noise models”. In: *Interspeech*. 2023, pp. 5102–5106. DOI: 10.21437/Interspeech.2023-232.

[C16] Xavier Alameda-Pineda, Qin Jin, Vincent Oria, and Laura Toni. “M4MM’22: 1st International Workshop on Methodologies for Multimedia”. In: *ACM International Conference on Multimedia*. Lisbon, Portugal, 2022. DOI: 10.1145/3503161.3554769.

[C17] Enrico Fini, Victor G Turrisi da Costa, Xavier Alameda-Pineda, Elisa Ricci, Karteek Alahari, and Julien Mairal. “Self-supervised models are continual learners”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022, pp. 9621–9630.

[C18] Laurent Girin, Xiaoyu Bie, Simon Leglaive, Thomas Hueber, and Xavier Alameda-Pineda. “Les auto-encodeurs variationnels dynamiques et leur application à la modélisation de spectrogrammes de parole”. In: *XXXIVe Journées d’Études sur la Parole*. 2022. DOI: 10.21437/JEP.2022-69.

[C19] Wen Guo, Xiaoyu Bie, Xavier Alameda-Pineda, and Francesc Moreno. “Multi-Person Extreme Motion Prediction with Cross-Interaction Attention”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022. DOI: 10.1109/CVPR52688.2022.01271.

[C20] Zhiqi Kang, Mostafa Sadeghi, Radu Horaud, Xavier Alameda-Pineda, Jacob Donley, and Anurag Kumar. “The impact of removing head movements on audio-visual speech enhancement”. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. 2022, pp. 7302–7306. DOI: 10.1109/ICASSP43922.2022.9746401.

[C21] Hanyu Xuan, Zhiliang Wu, Jian Yang, Yan Yan, and Xavier Alameda-Pineda. “A Proposal-based Paradigm for Self-supervised Sound Source Localization in Videos”. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022. DOI: 10.1109/CVPR52688.2022.00110.

[C22] Hanyu Xuan, Yihong Xu, Shuo Chen, Zhiliang Wu, Jian Yang, Yan Yan, and Xavier Alameda-Pineda. “Active Contrastive Set Mining for Robust Audio-Visual Instance Discrimination”. In: *International Joint Conference on Artificial Intelligence*. 2022. DOI: 10.24963/ijcai.2022/506.

[C23] Xiaoyu Bie, Laurent Girin, Simon Leglaive, Thomas Hueber, and Xavier Alameda-Pineda. “A Benchmark of Dynamical Variational Autoencoders applied to Speech Spectrogram Modeling”. In: *ISCA Interspeech*. 2021. DOI: 10.21437/Interspeech.2021-256.

[C24] Wen Guo, Enric Corona, Francesc Moreno-Noguer, and Xavier Alameda-Pineda. “PI-Net: Pose Interacting Network for Multi-Person Monocular 3D Pose Estimation”. In: *IEEE Winter Conference on Applications of Computer Vision*. 2021. DOI: 10.1109/WACV48630.2021.00284.

[C25] Viet-Nhat Nguyen, Mostafa Sadeghi, Elisa Ricci, and Xavier Alameda-Pineda. “Deep Variational Generative Models for Audio-visual Speech Separation”. In: *IEEE Workshop on Machine Learning for Signal Processing*. Queensland, Australia, 2021. DOI: 10.1109/MLSP52302.2021.9596406.

[C26] Mostafa Sadeghi and Xavier Alameda-Pineda. “Switching Variational Auto-Encoders for Noise-Agnostic Audio-visual Speech Enhancement”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. 2021. DOI: 10.1109/ICASSP39728.2021.9414097.

[C27] Xavier Alameda-Pineda, Miriam Redi, Jahna Otterbacher, Nicu Sebe, and Shih-Fu Chang. “FATE/MM’20: 2nd International Workshop on Fairness, Accountability, Transparency and Ethics”. In: *ACM International Conference on Multimedia*. 2020. DOI: 10.1145/3394171.3421896.

[C28] Guillaume Delorme, Yihong Xu, Stephane Lathuilière, Radu Horaud, and Xavier Alameda-Pineda. “CANU-ReID: A Conditional Adversarial Network for Unsupervised person Re-IDentification”. In: *IEEE International Conference on Pattern Recognition*. 2020. DOI: 10.1109/ICPR48806.2021.9412431.

[C29] Simon Leglaive, Xavier Alameda-Pineda, Laurent Girin, and Radu Horaud. “A Recurrent Variational Autoencoder for Speech Enhancement”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. 2020. DOI: 10.1109/ICASSP40776.2020.9053164.

[C30] Yahui Liu, Marco De Nadai, Deng Cai, Huayang Li, Xavier Alameda-Pineda, Nicu Sebe, and Bruno Lepri. “Describe What to Change: A Text-guided Unsupervised Image-to-Image Translation Approach”. In: *ACM International Conference on Multimedia*. 2020. DOI: 10.1145/3394171.3413505.

[C31] Mostafa Sadeghi and Xavier Alameda-Pineda. “Robust Unsupervised Audio-visual Speech Enhancement Using a Mixture of Variational Autoencoders”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. Barcelona, Spain, 2020. DOI: 10.1109/ICASSP40776.2020.9053730.

[C32] Yihong Xu, Aljosa Osep, Yutong Ban, Radu Horaud, Laura Leal-Taixé, and Xavier Alameda-Pineda. “How to Train Your Deep Multi-Object Tracker”. In: *IEEE International Conference on Computer Vision and Pattern Recognition*. Seatle,USA, 2020. DOI: 10.1109/CVPR42600.2020.00682.

[C33] Xavier Alameda-Pineda, Soraya Arias, Yutong Ban, Guillaume Delorme, Laurent Girin, Radu Horaud, Xiaofei Li, Bastien Mourgue, and Guillaume Sarrazin. “Audio-Visual Variational Fusion for Multi-Person Tracking with Robots”. In: *ACM Multimedia*. Nice, France, 2019. DOI: 10.1145/3343031.3350590.

[C34] Xavier Alameda-Pineda, Miriam Redi, Elisa Celis, Nicu Sebe, and Shih-Fu Chang. “FAT/MM’19: 1st International Workshop on Fairness, Accountability, and Transparency in MultiMedia”. In: *ACM International Conference on Multimedia*. Nice, France, 2019. DOI: 10.1145/3343031.3350555.

[C35] Xavier Alameda-Pineda, Miriam Redi, Nicu Sebe, Shih-Fu Chang, and Jiebo Luo. “ACM MM’18 Workshop on Understanding Subjective Attributes of Data, Multimodal Recognition of Evoked Emotions”. In: *ACM International Conference on Multimedia*. Seoul, Korea, 2018. DOI: 10.1145/3240508.3243721.

[C36] Yutong Ban, Xiaofei Li, Xavier Alameda-Pineda, Laurent Girin, and Radu Horaud. “Accounting for Room Acoustics in Audio-Visual Multi-Speaker Tracking”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. 2018. DOI: 10.1109/ICASSP.2018.8462100.

[C37] Stéphane Lathuilière, Pablo Mesejo, Xavier Alameda-Pineda, and Radu Horaud. “DeepGUM: Learning Deep Robust Regression with a Gaussian-Uniform Mixture Model”. In: *European Conference on Computer Vision*. Munich, Germany, 2018. DOI: 10.1007/978-3-030-01228-1_13.

[C38] Wang Wei, Xavier Alameda-Pineda, Dan Xu, Elisa Ricci, and Nicu Sebe. “Every Smile is Unique: Landmark-Guided Diverse Smile Generation”. In: *IEEE International Conference on Computer Vision and Pattern Recognition*. Salt Lake City, USA, 2018. DOI: 10.1109/CVPR.2018.00740.

[C39] Xavier Alameda-Pineda, Andrea Pilzer, Dan Xu, Nicu Sebe, and Elisa Ricci. “Viraliency: Pooling local virality”. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017, pp. 6080–6088.

[C40] Xavier Alameda-Pineda, Miriam Redi, Mohammad Soleymani, Nicu Sebe, Shih-Fu Chang, and Samuel Gosling. “MUSA2 – First ACM Workshop on Multimodal Understanding of Social, Affective and Subjective Attributes”. In: *ACM Multimedia*. Mountain View, USA, 2017. DOI: 10.1145/3123266.3132057.

[C41] Yutong Ban, Xavier Alameda-Pineda, Fabien Badeig, Sileye Ba, and Radu Horaud. “Tracking a Varying Number of People with a Visually-Controlled Robotic Head”. In: *Intelligent Robots and Systems*. Vancouver,Canada, 2017. DOI: 10.1109/IROS.2017.8206274.

[C42] Laurent Girin, Thomas Hueber, and Xavier Alameda-Pineda. “Adaptation of a Gaussian Mixture Regressor to a New Input Distribution: Extending the C-GMR Framework”. In: *International Conference on Latent Variable Analysis and Signal Separation*. Grenoble, France, 2017. DOI: 10.1007/978-3-319-53547-0_43.

[C43] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Sharon Gannot, and Radu Horaud. “An EM algorithm for joint source separation and diarisation of multichannel convolutive mixtures”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. New Orleans, USA, 2017. DOI: 10.1109/ICASSP.2017.7951789.

[C44] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Radu Horaud, and Sharon Gannot. “Exploting the Intermittency of Speech for Joint Separation and Diarization”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. New Paltz, USA, 2017. DOI: 10.1109/WASPAA.2017.8169991.

[C45] Aliaksandr Siarohin, Gloria Zen, Cveta Majtanovic, Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. “How to Make an Image More Memorable? A Deep Style Transfer Approach”. In: *ACM International Conference on Multimedia Retrieval*. Bucharest, Romania, 2017. DOI: 10.1145/3078971.3078986.

[C46] Dan Xu, Wanli Ouyang, Xavier Alameda-Pineda, Elisa Ricci, Xiaogang Wang, and Nicu Sebe. “Learning Deep Structured Multi-Scale Features using Attention-Gated CRFs for Contour Prediction”. In: *Advances in Neural Information Processing Systems*. Long Beach, USA, 2017.

[C47] Xavier Alameda-Pineda, Elisa Ricci, Yan Yan, and Nicu Sebe. “Recognizing Emotions from Abstract Paintings using Non-Linear Matrix Completion”. In: *IEEE International Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA, 2016. DOI: 10.1109/CVPR.2016.566.

[C48] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Sharon Gannot, and Radu Horaud. “An inverse-gama source variance prior with factorized parametrization for audio source separation”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. Shangai, China, 2016, pp. 136–140. DOI: 10.1109/ICASSP.2016.7471652.

[C49] Sergey Tulyakov, Xavier Alameda-Pineda, Elisa Ricci, Lijun Yin, Jeffrey F. Cohn, and Nicu Sebe. “Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions”. In: *IEEE International Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA, 2016. DOI: 10.1109/CVPR.2016.263.

[C50] Wei Wang, Yan Yan, Feiping Nie, Xavier Alameda-Pineda, Shuicheng Yan, and Nicu Sebe. “Projective Unsupervised Flexible Embedding with Optimal Graph”. In: *British Machine Vision Conference*. York, United Kingdom, 2016.

[C51] Dan Xu, Xavier Alameda-Pineda, Jingkuan Song, Elisa Ricci, and Nicu Sebe. “Academic Coupled Dictionary Learning for Sketch-based Image Retrieval”. In: *ACM International Conference on Multimedia*. Amsterdam, The Netherlands, 2016. DOI: 10.1145/2964284.2964329.

[C52] Dan Xu, Jingkuan Song, Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. “Multi-Paced Dictionary Learning for Cross-Domain Retrieval and Recognition”. In: *IEEE International Conference on Pattern Recognition*. Cancun, Mexico, 2016. DOI: 10.1109/ICPR.2016.7900132.

[C53] Xavier Alameda-Pineda, Yan Yan, Elisa Ricci, Oswald Lanz, and Nicu Sebe. “Analyzing Free-standing Conversational Groups: A Multimodal Approach”. In: *ACM International Conference on Multimedia*. Brisbane, Australia, 2015, pp. 4–15. DOI: 10.1145/2733373.2806238.

[C54] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Sharon Gannot, and Radu Horaud. “A Variational EM Algorithm for the Separation of Moving Sound Sources”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. New Paltz, USA, 2015, pp. 1–5. DOI: 10.1109/WASPAA.2015.7336936.

[C55] Israel-Dejene Gebru, Xavier Alameda-Pineda, Radu Horaud, and Florence Forbes. “Audio-Visual Speaker Localization via Weighted Clustering”. In: *IEEE Workshop on Machine Learning for Signal Processing*. Reims, France, 2014, pp. 1–6. DOI: 10.1109/MLSP.2014.6958874.

[C56] Maxime Janvier, Xavier Alameda-Pineda, Laurent Girin, and Radu Horaud. “Sound Representation and Classification Benchmark for Domestic Robots”. In: *IEEE International Conference on Robotics and Automation*. Hong Kong, China, 2014, pp. 6285–6292. DOI: 10.1109/ICRA.2014.6907786.

[C57] Xavier Alameda-Pineda, Radu Horaud, and Bernard Mourrain. “The Geometry of Sound Source Localization Using Non-Coplanar Microphone Arrays”. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. New Paltz, USA, 2013, pp. 1–4. DOI: 10.1109/WASPAA.2013.6701849.

[C58] Xavier Alameda-Pineda, Jordi Sanchez-Riera, and Radu Horaud. “Benchmarking methods for audio-visual recognition using tiny training sets”. In: *IEEE International Conference on Acoustics, Speech and Signal Processing*. 2013. DOI: 10.1109/ICASSP.2013.6638341.

[C59] Jan Cech, Ravi Mittal, Antoine Deleforge, Jordi Sanchez-Riera, Xavier Alameda-Pineda, and Radu Horaud. “Active-Speaker Detection and Localization with Microphones and Cameras Embedded into a Robotic Head”. In: *IEEE-RAS International Conference on Humanoid Robots*. 2013. DOI: 10.1109/HUMANOIDS.2013.7029977.

[C60] Xavier Alameda-Pineda and Radu Horaud. “Geometrically-constrained Robust Time Delay Estimation Using Non-coplanar Microphone Arrays”. In: *European Signal Processing Conference*. Bucarest, Romania, 2012, pp. 1309–1313.

[C61] Maxime Janvier, Xavier Alameda-Pineda, Laurent Girin, and Radu Horaud. “Sound-event recognition with a companion humanoid”. In: *IEEE-RAS International Conference on Humanoid Robots*. 2012. DOI: 10.1109/HUMANOIDS.2012.6651506.

[C62] Jordi Sanchez-Riera, Xavier Alameda-Pineda, and Radu Horaud. “Audio-visual robot command recognition: D-META’12 grand challenge”. In: *ACM International Conference on Multimodal Interaction*. 2012. DOI: 10.1145/2388676.2388760.

[C63] Jordi Sanchez-Riera, Xavier Alameda-Pineda, Johannes Wienke, Antoine Deleforge, Soraya Arias, Jan Cech, Sebastian Wrede, and Radu Horaud. “Online Multimodal Speaker Detection for Humanoid Robots”. In: *IEEE-RAS International Conference on Humanoid Robotics*. Osaka, Japan, 2012, pp. 126–133. DOI: 10.1109/HUMANOIDS.2012.6651509.

[C64] Xavier Alameda-Pineda, Vasil Khalidov, Radu Horaud, and Florence Forbes. “Finding audio-visual events in informal social gatherings”. In: *IEEE/ACM International Conference on Multimodal Interaction*. 2011. DOI: 10.1145/2070481.2070527.

[C65] Julio Rolón, Philippe Salembier, and Xavier Alameda-Pineda. “Image compression with generalized lifting and partial knowledge of the signal pdf”. In: *IEEE International Conference on Image Processing*. 2008. DOI: 10.1109/ICIP.2008.4711708.

Workshop Papers

[W1] Thomas De Min, Massimiliano Mancini, Karteek Alahari, Xavier Alameda-Pineda, and Elisa Ricci. “On the Effectiveness of LayerNorm Tuning for Continual Learning in Vision Transformers”. In: *International Conference on Computer Vision Workshops*. 2023.

[W2] David Emukpere, Xavier Alameda-Pineda, and Chris Reinke. “Successor Feature Neural Episodic Control”. In: *Fifth Workshop on Meta-Learning at the Conference on Neural Information Processing Systems*. 2021.

[W3] Guillaume Delorme, Yutong Ban, Guillaume Sarrazin, and Xavier Alameda-Pineda. “ODA-Track: Online Deep Appearance for Robotic Multiple Person Tracking”. In: *IAPR International Conference on Pattern Recognition Workshops*. 2020.

[W4] Mihai Gabriel Constantin, Bogdan Ionescu, Claire-Hélène Demarty, Ngoc QK Duong, Xavier Alameda-Pineda, and Mats Sjöberg. “Predicting Media Memorability Task at MediaEval 2019”. In: *MediaEval 2019 Workshop*. 2019.

[W5] X. Li, Y. Ban, L. Girin, X. Alameda-Pineda, and R. Horaud. “A cascaded multiple-speaker localization and tracking system”. In: *International Workshop on Acoustic Signal Enhancement (IWAENC), LOCATA Satellite Workshop*. Tokyo, Japan, 2018.

[W6] Yutong Ban, Laurent Girin, Xavier Alameda-Pineda, and Radu Horaud. “Exploiting the Complementarity of Audio-Visual Data for Probabilistic Multi-Speaker Tracking”. In: *IEEE ICCV Workshop on Computer Vision for Audio-Visual Media*. Venice, Italy, 2017. DOI: 10.1109/ICCVW.2017.60.

[W7] Yutong Ban, Sileye Ba, Xavier Alameda-Pineda, and Radu Horaud. “Tracking Multiple Persons Based on a Variational Bayesian Model”. In: *European Conference on Computer Vision Workshops*. Amsterdam, 2016, pp. 52–67.

Thesis and Dissertations

[TH1] Xavier Alameda-Pineda. “Towards Probabilistic Generative Models for Socially Intelligent Robot”. Habilitation à Diriger de Recherches. PhD thesis. Université Grenoble Alpes, 2020.

[TH2] Xavier Alameda-Pineda. “Egocentric Audio-Visual Scene Analysis, A Machine Learning and Signal Processing Approach”. PhD thesis. Université Joseph Fourier, Grenoble, France, 2013.

[TH3] Xavier Alameda-Pineda. “Finding audio-visual events with a robot head”. MA thesis. Université Joseph Fourier / ENSIMAG, Grenoble, France, 2010.

[TH4] Xavier Alameda-Pineda. “Entropy Coding for Image Compression Based on Generalized Lifting and SPECK”. MA thesis. ETSETB, BarcelonaTECH, Barcelona, Spain, 2009.

Books

[B1] Xavier Alameda-Pineda, Elisa Ricci, Nicu Sebe, et al. *Multimodal Behavior Analysis in the Wild: Advances and Challenges*. Ed. by Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. Elsevier, 2018. DOI: 10.1016/C2017-0-01387-3.

Book Chapters

[CH1] Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. “Multimodal behavior analysis in the wild: an introduction”. In: *Multimodal Behavior Analysis in the Wild*. Ed. by Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. Elsevier, 2018, pp. 1–10. DOI: 10.1016/B978-0-12-814601-9.00011-0.

[CH2] Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. “Multimodal analysis of free-standing conversational groups”. In: *Frontiers of Multimedia Research*. Ed. by Shih-Fu Chang. Morgan and Claypool, 2017, pp. 51–74. DOI: 10.1145/3122865.3122869.

[CH3] Xavier Alameda-Pineda, Ramanathan Subramanian, Elisa Ricci, Oswald Lanz, and Nicu Sebe. “SALSA: A multimodal dataset for the automated analysis of free-standing social interactions”. In: *Group and Crowd Behavior for Computer Vision*. Ed. by Vittorio Murino, Marco Cristani, Shishir Shah, and Silvio Savarese. Elsevier, 2016. DOI: 10.1016/B978-0-12-809276-7.00017-5.