

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

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

xavier.alameda-pineda@inria.fr, @xavirema

(last updated on April 2025)



## 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 (equivalent) in Telecommunications Engineering [TH4]** (ETSETB - BarcelonaTECH, Barcelona, Spain)
- 2008 – **Masters (equivalent) 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 – previously **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 in the STC at UPC, Barcelona, Spain)
- 2007-2008 – **Research collaboration grant for young researchers** [WITH PROF. XAVIER CABRÉ] (MEC and MA1 at UPC, Barcelona, Spain)

## AWARDS AND HONORS

- IEEE TMM 2022 Outstanding Associate Editor Award** for contributions to the IEEE TMM Editorial Board at [RobotLearn Web Site](#)
- ACM TOMM 2020 Nicolas D. Georganas Best Paper Award** for “Increasing image memorability with neural style transfer” [J25] at [ACM Web Site](#)
- ACM SIGMM 2018 Rising Star Award** for Contributions to Multimodal Social Behavior Understanding at [SIGMM Web Site](#)
- IEEE IROS 2017 Novel Technology Award Finalist** for Tracking a Varying Number of People with a Visually-Controlled Robotic Head [C38] at [International Conference on Intelligent Robots and Systems’17](#)
- IAPR ICPR 2016 Best scientific paper award** for Multi-Paced Dictionary Learning for Cross-Domain Retrieval and Recognition [C49] at [IAPR International Conference on Pattern Recognition’16](#)
- ACM MM 2015 Best paper award** for Analysing Free-standing Conversational Groups: A Multimodal Approach [C50] at [ACM International Conference on Multimedia’15](#)
- IEEE WASPAA 2015 Best student paper award** for A Variational EM Algorithm for the Separation of Moving Sound Sources [C51] at [IEEE Workshop on Applications of Signal Processing to Audio and Acoustics’15](#)
- ACM ICMI 2011 Outstanding paper award** for Finding Audio-visual Events in Informal Social Gatherings [C61] at [IEEE/ACM International Conference on Multimodal Interaction’11](#)

## 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
<b>PhD (co-)Supervision – (co-)Directeur de thèse</b>			
<b>Jordan Cosio</b>	2023- (Inria Grenoble)	Ongoing [C6].	Dr. Pierre-Brice Wieber
<b>Jean-Eudes Ayilo</b>	2023- (Inria Nancy)	Ongoing [C1].	Dr. Mostafa Sadeghi & Prof. Romain Serizel
<b>Gaétan Lepage</b>	2020- (Inria Grenoble)	Ongoing	Prof. Laurent Girin & Dr. Chris Reinke
<b>Guérolé Fiche</b>	2023-24 (Central Supélec)	Defended (18/11/2024) [C8, C4, C2]	Prof. Simon Leglaive
<b>Samir Sadok</b>	2023-24 (Central Supélec)	Defended (08/03/2024) [J9, J3]	Prof. Simon Leglaive
<b>Xiaoyu Lin</b>	2020-24 (Inria Grenoble)	Defended (25/06/2024) [C11, J7, C12]	Prof. Laurent Girin
<b>Anand Ballou</b>	2019-24 (Inria Grenoble)	Defended (27/03/2024) [J5]	Dr. Chris Reinke
<b>Louis Airale</b>	2019- (Inria Grenoble)	Defended (04/12/2023) [J10, J1]	Prof. Dominique Vaufreydaz
<b>Xiaoyu Bie</b>	2019- (Inria Grenoble)	Defended (20/10/2023) [C20, J12, J11, C16]	Prof. Laurent Girin
<b>Wen Guo</b>	2019-2023 (Inria Grenoble)	Defended (12/06/2023) [C21, C16, C10]	Dr. Francesc Moreno-Noguer
<b>Guanglei Yang</b>	2021-22 (HIT)	Defended (12/06/2023) [J14, J15]	Prof. Elisa Ricci
<b>Enrico Fini</b>	2021-22 (U. Trento)	Defended (17/04/2023) [C14, C9]	Prof. Elisa Ricci & Dr. Karteek Alahari
<b>Hanyu Xuan</b>	2021-22 (NUST)	Defended (2022) [C18, C19]	Prof. Yan Yan
<b>Yihong Xu</b>	2018-22 (Inria Grenoble)	Defended (08/06/2022) [C29, J13, C19, C25]	Dr. Radu Horaud
<b>Guillaume Delorme</b>	2017-21 (Inria Grenoble)	Defended (08/10/2021) [C25, W3, J13, C30]	Dr. Radu Horaud
<b>Aliaksandr Siarohin</b>	2016-17 (U. Trento)	Defended (24/06/2021) [C42, J25]	Prof. Nicu Sebe
<b>Andrea Pilzer</b>	2016-17 (U. Trento)	Defended (22/06/2020) [C36]	Prof. Nicu Sebe
<b>Yutong Ban</b>	2016-19 (Inria Grenoble)	Defended (10/05/2019) [W6, W7, C38, J22, J17, C33]	Dr. Radu Horaud
<b>Wei Wang</b>	2016-18 (U. Trento)	Defended (2018) [C47, C35, J20]	Prof. Nicu Sebe
<b>Stéphane Lathuilière</b>	2016-17 (Inria Grenoble)	Defended (22/05/2018) [J23, C34, C25]	Dr. Radu Horaud
<b>Israel-Dejene Gebru</b>	2014-17 (Inria Grenoble)	Defended (13/04/2018) [C52, J31]	Dr. Radu Horaud
<b>Dan Xu</b>	2014-17 (U. Trento)	Defended (2018) [C43, J21, J26, C48, C49]	Prof. Nicu Sebe
<b>Sergey Tulyakov</b>	2015-17 (U. Trento)	Defended (2017) [C46]	Prof. Nicu Sebe
<b>Vicent Drouard</b>	2016-17 (Inria Grenoble)	Defended (18/12/2017) [J16]	Dr. Radu Horaud
<b>Dionyssos Kounades-Bastian</b>	2014-17 (Inria Grenoble)	Defended (24/02/2017) [J32, C45, C40, C51, C41]	Dr. Radu Horaud & Prof. Laurent Girin
<b>MSc (co-)Supervision – (co-)Encadrant de Stage</b>			
<b>Maxime Attwood</b>	Defended (2024) (Inria Grenoble)	2025	Dr. Samir Sadok
<b>Ahmad Ghazi</b>	Defended (2024) (Inria Grenoble)	2024	Dr. Pedro Rodrigues
<b>David Emukpere</b>	2021 (Inria Grenoble)	Defended (2021) [W2]	Dr. Chris Reinke
<b>Viet Nhat Nguyen</b>	2020 (Inria Grenoble)	Defended (2020) [C22]	Dr. Mostafa Sadeghi
<b>Álvaro González Jiménez</b>	2020 (Inria Grenoble)	Defended (2020)	Prof. Stéphane Lathuilière
<b>Predrag Pilipovic</b>	2020 (Inria Grenoble)	Defended (2020)	
<b>Vadim Sushko</b>	2019 (Inria Grenoble)	Defended (2019)	
<b>Elsa Marie</b>	2019 (Inria Grenoble)	Defended (2019)	Prof. Simon Leglaive
<b>Divya Grover</b>	2018 (Inria Grenoble)	Defended (2018)	Dr. Radu Horaud
<b>Alessio Xompero</b>	2014 (Inria Grenoble)	Defended (2014)	Dr. Radu Horaud
<b>Israel-Dejene Gebru</b>	2013 (Inria Grenoble)	Defended (2013)	Dr. Radu Horaud
<b>Maxime Janvier</b>	2012 (Inria Grenoble)	Defended (2012)	Dr. Radu Horaud

## TEACHING

2024-2025 – <b>Advanced Machine Learning: Applications to Vision, Audio and Text (12h)</b>	[WITH DR. KARTEEK ALAHARI, DR. ERIC GAUSSIER]	(ENSIMAG/UGA, Grenoble, France)
2023-2024 – <b>Learning, Probabilities and Causality (15h)</b>	[WITH DR. THOMAS HUEBER, DR. ERIC GAUSSIER]	(ENSIMAG/UGA, Grenoble, France)
2023-2024 – <b>Advanced Machine Learning: Applications to Vision, Audio and Text (12h)</b>	[WITH DR. KARTEEK ALAHARI, DR. ERIC GAUSSIER]	(ENSIMAG/UGA, Grenoble, France)
2022-2023 – <b>Learning, Probabilities and Causality (15h)</b>	[WITH DR. THOMAS HUEBER, DR. ERIC GAUSSIER]	(ENSIMAG/UGA, Grenoble, France)
2022-2023 – <b>Advanced Machine Learning: Applications to Vision, Audio and Text (12h)</b>	[WITH DR. KARTEEK ALAHARI, DR. ERIC GAUSSIER]	(ENSIMAG/UGA, Grenoble, France)
2021-2022 – <b>Fundamentals of Probabilistic Data Mining (15h)</b>	[WITH DR. THOMAS HUEBER]	(ENSIMAG/UGA, Grenoble, France)
2021-2022 – <b>Machine Learning for Multimodal Data (12h)</b>	[WITH DR. KARTEEK ALAHARI, DR. ERIC GAUSSIER]	(ENSIMAG/UGA, Grenoble, France)
2020-2021 – <b>Fundamentals of Probabilistic Data Mining (15h)</b>	[WITH DR. THOMAS HUEBER]	(ENSIMAG/UGA, Grenoble, France)
2020-2021 – <b>Machine Learning for Computer Vision and Audio Processing (12h)</b>	[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 Proc. (2022-2024), IAPR TC9 on Pattern Recog. in Human Machine Interaction (2021-2023).

**Award Chair:** IEEE ICME 2024 Best Paper Award Committee co-Chair, IEEE TMM 2023 Best Paper Award Committee co-Chair

**Associated Editor:** (period, # handled papers) ACM TOMM (2019-, 41), IEEE TMM (2021-2024, 55), CVIU (2021-, 14), ACM TIST (2021-2024, 11).

**Special Issue Guest Editor:** [Eyes on People: Recent Trends on Human Analysis, Perception and Generation](#) (ongoing, Computer Vision and Image Understanding), [Multi-Modal Understanding of Social, Affective and Subjective Attributes of Data](#) (4 papers, 2019, ACM Transactions on Multimedia Computing, Communications, and Applications), [Generating Realistic Visual Data of Human Behavior](#) (8 papers, 2020, International Journal of Computer Vision).

**General Chair:** ACM MM 2026.

**Program Chair:** ACM MM 2022.

**Senior Area Chair** (# handled papers): ACM MM 2025, ACM MM 2024 (185).

**Area Chair** (# handled papers): ICRA'24 (11), ACM MM'23 (25), IEEE WASPAA'23 (4), IEEE ICASSP'23 (12), ICRA'23 (11), [BMVC'22](#) (8), AAAI'22 (2), ACM MM'21 (9), IEEE WACV'21 (20), ACM MM'20 (26), IAPR ICPR'20 (20), [ACM MM'19](#) (24), [ICIA'19](#) (17), [ICCV'17](#) (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'19](#), [WHBU](#) @ ICCV 2019, [FAT-MM](#) @ ACM MM 2019, [FFSS-USAD](#) @ IEEE/CVF CVPR 2019, [HBUGEN](#) @ ECCV 2018, [EE-USAD](#) @ ACM MM 2018, [V-USAD](#) @ IEEE/CVF CVPR 2018.

**Special Session Organiser:** IEEE ISCAS'20, IEEE ICIP'19, ACM ICMR'17

**Reviewing for International Journals** (# reviewed papers) – Not updated since 2023: IEEE Trans. Multimedia (19) IEEE/ACM Trans. Audio, Signal, and Language Processing (13), IEEE Trans. Pattern Analysis and Machine Intelligence (7), IEEE Trans. Image Processing (3), IEEE Trans. Signal Processing (4), IEEE Trans. Circuits and Systems for Video Technology (5), Computer Vision and Image Understanding (2).

**Reviewing for International Conferences** – Not updated since 2023: IEEE CVPR'17-19 & 23, NIPS'16-19, ICLR'18-19, ECCV'16,18, ACM MM'16-19, ICM'12-16.

**PhD Reviewing:** [Hakim Benkirane](#) (U. Paris Saclay'24), [Remi Rigal](#) (ENSTA Bretagne'23), [Stefanie Tan](#) (TU Delf'23), [Jean-Yves Franceschi](#) (Sorbonne Universités'22), [Manuel Pariente](#) (U. Lorraine'21), [Marco Godi](#) (U. Verona'21), [Daniel Michelsanti](#) (Aalborg University'20), [Wei Wang](#) (University of Trento'18), [Dan Xu](#) (University of Trento'18), [Daniek Brink](#) (University of Stellenbosch'16), [José Velasco](#) (University of Alcalá'16).

**PhD Committee:** [Ilyass Moummad](#) (IMT Atlantique'24), [Ruben Delgado Escaño](#) (U. Málaga'22), [Julien Audibert](#) (U. Sorbonne'21), [Maria Kabtoul](#) (University Grenoble-Alpes'21), [Stéphane Lathuilière](#) (University Grenoble-Alpes'18), [Irtiza Hasan](#) (University of Verona'19), [Theodoros Tsesmelis](#) (University of Verona'19), [Israel-Dejene Gebru](#) (University of Grenoble-Alpes'18), [Vincent Drouard](#) (University of Grenoble Alpes'17), [Dionyssos Kounades-Bastian](#) (University of Grenoble Alpes'17).

**Misc:** SIGMM Conference Ambassadors Program'19 for fairness, accountability and transparency brought to multimedia

## TUTORIALS & INVITED TALKS

**Probabilistic generative models for audio-visual processing – invited course** (Oct'24)

at [Paris GenAI Autumn School](#), Saclay, FR

**Learning for Companion Robots: Preparation and Adaptation – keynote talk** (July'24)

at [Joint RFIAP-cAP](#), Lille, FR – [Slides](#)

**Social Robot Learning – invited talk** (June'24)

at CEA List Days, Saclay, FR

**Variational Audio-Visual Representation Learning – keynote talk** (Nov'23)

at [ACM International Conference on Multimedia](#), Ottawa, CA – [Slides](#)

**Learning for Robots in Conversational Groups – invited talk** (May'23)

at [Workshop of the International Laboratory on Learning Systems](#), Université Paris Saclay

**Robots within Groups of People – invited talk** (May'23)

at Interdisciplinary Workshop on Mingling Technologies, TU Delft

**Unsupervised Probabilistic Learning with Latent Variables – invited course** (Jan'23)

at [Machine Learning Summer School Africa 2023](#) – [Slides](#)

**Learning for Socially Intelligent Robots – invited talk** (Dec'22)

at Computer Science and Electric Engineering Departments, University of Alberta

**Introduction to Dynamical Variational Autoencoders – invited talk** (Feb'22)

at [MLIA Research Team Seminars](#)

**Deep Generative Modeling of Sequential Data with Dynamical Variational Autoencoders – tutorial** (*Jun'21*)  
at [IEEE International Conference on Audio, Speech and Signal Processing – Recording](#)

**Unsupervised Learning for Human Robot Perception – invited talk** (*Jun'21*)  
at [Robotics and AI Summer School 2021 – Recording](#)

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

**Unsupervised Audio-Visual Fusion for Upstream Human Behavior Understanding – invited talk** (*May'21*)  
at [AI4Media Workshop on New Learning Paradigms and Distributed AI – Recording](#) (starts at 1h09')

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

**Speaker localisation and enhancement in populated environments – invited talk** (*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 talk** (*Jan'21*)  
at [ICPR 2020 Workshop on Multimodal pattern recognition for social signal processing in human computer interaction](#)

**Towards audio-visual speech enhancement in robotic platforms – invited talk** (*Dec'20*)  
at [Journée "perception et interaction homme-robot" du Groupe de Travail GT5 Interactions Personnes / Systèmes Robotiques du GDR Robotique](#)

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

**Choosing wisely your deep training loss – invited talk** (*March'20*)  
at [Universidade NOVA de Lisboa](#)

**Artificial Intelligence for Social Robots in Gerontological Healthcare – invited talk** (*March'20*)  
at [European Robotics Forum](#)

**Probabilistic and deep learning for regression in computer vision – tutorial** (*Sep'19*)  
at [International Conference on Image Analysis and Processing](#)

**Significance & Robustness in Deep Regression – invited talk** (*Jul'19*)  
at [University of Trento](#)

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

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

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

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

**Multimodal human behavior analysis in the wild – tutorial** (*Dec'16*)  
at [IAPR International Conference on Pattern Recognition](#)

**Emerging topics in noisy and missing data – tutorial** (*Oct'16*)  
at [ACM International Conference on Multimedia](#)

**Matrix completion: a computer vision perspective – seminar** (*Jun'16*)  
at [Carnegie Mellon University](#) and [Digital Video and Multimedia Lab of Columbia University](#)

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

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

**Free-standing conversational groups: the SALSA dataset and multi-modal head and body pose estimation – invited talk** (*Nov'15*)  
at [Universitat Politècnica de Catalunya - Image and Video Processing Group](#) and [Inria Nancy Grand-Est - Team multispeech](#)

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

## SOFTWARE

**AnCoGen** (Samir Sadok, PhD Student, 2025): Analysis, Control and Generation of Speech with a Masked Autoencoder [C3].

**MEGA** (Guéno   Fich  , PhD Student, 2025): Masked Generative Autoencoder for Human Mesh Recovery [C2].

**AV-UDiffuse+** (Jean-Eudes Ayilo, PhD Student, 2025): Diffusion-based Unsupervised Audio-visual Speech Enhancement [C1].

**Lost&Found** (Lorenzo Vaquero Otal, PhD Student, 2024): Overcoming Detector Failures in Online Multi-Object Tracking [C7].

**VQ-HPS** (Gu  no   Fich  , PhD Student, 2024): Vector-Quantized Human Pose and Shape Estimation [C4].

**MDVAE** (Samir Sadok, PhD Student, 2024): Multimodal dynamical variational autoencoder [J3].

**SUMHoE** (Louis Airale, PhD Student, 2023): Semantic unconditional head motion generation [J1].

**Motion-DVAE** (Guénolé Fiche, PhD Student, 2023): Fast Human motion denoising [C8].

**DDGM-SE** (Xiaoyu Lin, PhD Student, 2023): deep diffusion generative model for speech enhancement [C12].

**MixDVAE** (Xiaoyu Lin, PhD Student, 2023): unsupervised multiple source tracking/separation library using DVAE [J7].

**SFR Learning** (Chris Reinke, SRP, 2022): implementation of our generalized transfer reinforcement learning [J8].

**LigHT-DVAE** (Xiaoyu Lin, PhD Student, 2023): code of the light version of the HIT-DVAE, named LigHT-DVAE [C11].

**CASSLE** (Enrico Fini, PhD Student, 2022): code for the paper self-supervised models are continual learners [C14].

**UCD** (Guanglei Yang, PhD Student, 2022): uncertainty aware distillation for semantic segmentation [J15].

**Multi-party Interaction Simulator** (Chris Reinke, SRP and Alex Auteraud, Engineer, 2022): multi-party interaction simulator for RL-based training.

**TransCenter** (Yihong Xu, PhD Student, 2022): transformer-based architecture for multiple object center tracking [J13].

**CAF** (Guanglei Yang, PhD Student, 2022): continual attentive fusion for incremental semantic segmentation [J14].

**SFNEC** (David Emukpere, MSc Student, 2021): software for the successor feature neural episodic control [W2].

**PI-NET** (Wen Guo, PhD Student, 2021): software for multi-person monocular 3D pose estimation [C21].

**Social MPC and Robot Behavior** (Timothée Wintz, SRP, Alex Auteraud, Engineer, Chris Reinke, SRP, 2021-): software implementation of the socially-aware model predictive control framework working on the ARI robot.

**DVAE** (Xiaoyu Bie, PhD Student, 2020): training and evaluating dynamical variational autoencoder [J12, C20].

**DeepMOT** (Yihong Xu, PhD Student, 2020): a paradigm for training multiple object trackers [C29].

**CANU-REID** (Guillaume Delorme, PhD Student, 2019): adversarial unsupervised domain adaptation [C25].

**AVSE-VAE and MIN-VAE** (Mostafa Sadeghi, SRP, 2019 and 2020): VAE-based audio-visual speech enhancement [J19, C28, C23, J18].

**Deep Regression** (Stéphane Lathuilière, PhD Student, 2018): deep regression benchmark library of [J23].

**DeepGUM** (Stéphane Lathuilière, PhD Student, 2018): implementation of the deep Gaussian-uniform mixture [C34].

**EM-WD** (Israel Gebru, PhD Student, 2016): expectation-maximisation algorithm with weighted data [C52, J31].

**EMD and DnD** (Dionysos Kounades, PhD Student, 2017): EM algorithm for simultaneous sound source separation and diarisation [C40, C41].

**REMOVE** (Dionysos Kounades, PhD Student, 2016): variational EM for separating moving sources [J32].

**IGNMF** (Dionysos Kounades, PhD Student, 2016): EM algorithm for sound separation via inverse-gamma [C45].

## PUBLICATIONS

### JOURNAL ARTICLES

- [J1] 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](https://doi.org/10.1145/3635154).
- [J2] 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](https://doi.org/10.1016/j.neucom.2023.126941).
- [J3] 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).
- [J4] 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](https://doi.org/10.1109/TPAMI.2024.3363508).
- [J5] 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](https://doi.org/10.1007/s10489-023-04691-5).
- [J6] 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](https://doi.org/10.1007/s11263-022-01742-1).
- [J7] 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).
- [J8] Chris Reinke and Xavier Alameda-Pineda. "Successor Feature Representations". In: *Transactions on Machine Learning Research* (2023).
- [J9] 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](https://doi.org/10.1016/j.specom.2023.02.005).
- [J10] 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](https://doi.org/10.1109/TAFFC.2022.3171719).
- [J11] 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](https://doi.org/10.1109/TASLP.2022.3207349).
- [J12] 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: [10.1561/22000000089](https://doi.org/10.1561/22000000089).

- [J13] 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](https://doi.org/10.1109/TPAMI.2022.3225078).
- [J14] 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](https://doi.org/10.1109/TMM.2022.3167555).
- [J15] 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](https://doi.org/10.1109/TPAMI.2022.3163806).
- [J16] 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](https://doi.org/10.1109/TNNLS.2021.3054407).
- [J17] 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](https://doi.org/10.1109/TPAMI.2019.2953020).
- [J18] 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](https://doi.org/10.1109/TSP.2021.3066038).
- [J19] 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](https://doi.org/10.1109/TASLP.2020.3000593).
- [J20] 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](https://doi.org/10.1109/TMM.2019.2963621).
- [J21] 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](https://doi.org/10.1109/TPAMI.2020.3043781).
- [J22] 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](https://doi.org/10.1109/LSP.2019.2908376).
- [J23] Stéphane 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](https://doi.org/10.1109/TPAMI.2019.2910523).
- [J24] 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* 1 (2019), pp. 88–103. doi: [10.1109/JSTSP.2019.2903472](https://doi.org/10.1109/JSTSP.2019.2903472).
- [J25] 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](https://doi.org/10.1145/3311781).
- [J26] 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](https://doi.org/10.1109/TIP.2018.2837381).
- [J27] 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](https://doi.org/10.1016/j.specom.2017.08.002).
- [J28] 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](https://doi.org/10.1109/TASLP.2017.2651398).
- [J29] 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](https://doi.org/10.1109/TPAMI.2015.2496269).
- [J30] 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](https://doi.org/10.1016/j.cviu.2016.07.006).
- [J31] 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](https://doi.org/10.1109/TPAMI.2016.2522425).
- [J32] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Sharon Gannot, and Radu Horaud. "A Variational EM Algorithm for the Separation of Time-Varying Convolutional Audio Mixtures". In: *IEEE/ACM Transactions on Audio, Speech and Language Processing* 24.8 (2016), pp. 1408–1423. doi: [10.1109/TASLP.2016.2554286](https://doi.org/10.1109/TASLP.2016.2554286).
- [J33] 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](https://doi.org/10.1177/0278364914548050).
- [J34] 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](https://doi.org/10.1109/TASLP.2015.2464702).
- [J35] 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](https://doi.org/10.1109/TASLP.2014.2317989).
- [J36] 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](https://doi.org/10.1007/s12193-012-0111-y).

## CONFERENCE PAPERS

- [C1] 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.
- [C2] Guérolé 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.
- [C3] 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.

- [C4] 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 Computer Vision*. 2024. doi: [10.1007/978-3-031-72943-0\\_27](https://doi.org/10.1007/978-3-031-72943-0_27).
- [C5] 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](https://doi.org/10.1109/ICASSP48485.2024.10446294).
- [C6] 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.
- [C7] 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](https://doi.org/10.1007/978-3-031-73464-9_27).
- [C8] 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](https://doi.org/10.1145/3623264.3624454).
- [C9] 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](https://doi.org/10.1109/CVPR52729.2023.00311).
- [C10] 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](https://doi.org/10.1109/WACV56688.2023.00479).
- [C11] 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](https://doi.org/10.1109/ICASSP49357.2023.10096751).
- [C12] 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](https://doi.org/10.21437/Interspeech.2023-232).
- [C13] 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](https://doi.org/10.1145/3503161.3554769).
- [C14] Enrico Fini, Victor G Turrissi 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.
- [C15] 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](https://doi.org/10.21437/JEP.2022-69).
- [C16] 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](https://doi.org/10.1109/CVPR52688.2022.01271).
- [C17] 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](https://doi.org/10.1109/ICASSP43922.2022.9746401).
- [C18] 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](https://doi.org/10.1109/CVPR52688.2022.00110).
- [C19] 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](https://doi.org/10.24963/ijcai.2022/506).
- [C20] 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](https://doi.org/10.21437/Interspeech.2021-256).
- [C21] 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](https://doi.org/10.1109/WACV48630.2021.00284).
- [C22] 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](https://doi.org/10.1109/MLSP52302.2021.9596406).
- [C23] 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](https://doi.org/10.1109/ICASSP39728.2021.9414097).
- [C24] 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](https://doi.org/10.1145/3394171.3421896).
- [C25] Guillaume Delorme, Yihong Xu, Stéphane 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](https://doi.org/10.1109/ICPR48806.2021.9412431).
- [C26] 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](https://doi.org/10.1109/ICASSP40776.2020.9053164).
- [C27] 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](https://doi.org/10.1145/3394171.3413505).
- [C28] 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](https://doi.org/10.1109/ICASSP40776.2020.9053730).
- [C29] 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*. Seattle, USA, 2020. doi: [10.1109/CVPR42600.2020.00682](https://doi.org/10.1109/CVPR42600.2020.00682).
- [C30] 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](https://doi.org/10.1145/3343031.3350590).
- [C31] 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](https://doi.org/10.1145/3343031.3350555).
- [C32] 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](https://doi.org/10.1145/3240508.3243721).

- [C33] 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](https://doi.org/10.1109/ICASSP.2018.8462100).
- [C34] 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](https://doi.org/10.1007/978-3-030-01228-1_13).
- [C35] 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](https://doi.org/10.1109/CVPR.2018.00740).
- [C36] 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.
- [C37] 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](https://doi.org/10.1145/3123266.3132057).
- [C38] Yutong Ban, Xavier Alameda-Pineda, Fabien Badeig, Siley 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/IR0S.2017.8206274](https://doi.org/10.1109/IR0S.2017.8206274).
- [C39] 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](https://doi.org/10.1007/978-3-319-53547-0_43).
- [C40] 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](https://doi.org/10.1109/ICASSP.2017.7951789).
- [C41] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Radu Horaud, and Sharon Gannot. "Exploiting 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](https://doi.org/10.1109/WASPAA.2017.8169991).
- [C42] 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](https://doi.org/10.1145/3078971.3078986).
- [C43] 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.
- [C44] 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](https://doi.org/10.1109/CVPR.2016.566).
- [C45] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Sharon Gannot, and Radu Horaud. "An inverse-gamma source variance prior with factorized parametrization for audio source separation". In: *IEEE International Conference on Audio, Speech and Signal Processing*. Shanghai, China, 2016, pp. 136–140. doi: [10.1109/ICASSP.2016.7471652](https://doi.org/10.1109/ICASSP.2016.7471652).
- [C46] 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](https://doi.org/10.1109/CVPR.2016.263).
- [C47] 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.
- [C48] 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](https://doi.org/10.1145/2964284.2964329).
- [C49] 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](https://doi.org/10.1109/ICPR.2016.7900132).
- [C50] 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](https://doi.org/10.1145/2733373.2806238).
- [C51] 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](https://doi.org/10.1109/WASPAA.2015.7336936).
- [C52] 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](https://doi.org/10.1109/MLSP.2014.6958874).
- [C53] 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](https://doi.org/10.1109/ICRA.2014.6907786).
- [C54] 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](https://doi.org/10.1109/WASPAA.2013.6701849).
- [C55] 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*. Vancouver, Canada, 2013, pp. 3662–3666. doi: [10.1109/ICASSP.2013.6638341](https://doi.org/10.1109/ICASSP.2013.6638341).
- [C56] 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*. Atlanta, USA, 2013, pp. 203–210. doi: [10.1109/HUMANOIDS.2013.7029977](https://doi.org/10.1109/HUMANOIDS.2013.7029977).
- [C57] 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.
- [C58] Maxime Janvier, Xavier Alameda-Pineda, Laurent Girin, and Radu Horaud. "Sound-Event Recognition with a Companion Humanoid". In: *IEEE-RAS International Conference on Humanoid Robotics*. Osaka, Japan, 2012, pp. 104–111. doi: [10.1109/HUMANOIDS.2012.6651506](https://doi.org/10.1109/HUMANOIDS.2012.6651506).
- [C59] Jordi Sanchez-Riera, Xavier Alameda-Pineda, and Radu Horaud. "Audio-Visual Robot Command Recognition". In: *IEEE/ACM International Conference on Multimodal Interaction*. Santa Monica, USA, 2012, pp. 371–378. doi: [10.1145/2388676.2388760](https://doi.org/10.1145/2388676.2388760).



- [C60] 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](https://doi.org/10.1109/HUMANOIDS.2012.6651509).
- [C61] 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 Interfaces*. Alicante, Spain, 2011, pp. 247–254. doi: [10.1145/2070481.2070527](https://doi.org/10.1145/2070481.2070527).
- [C62] Julio C. Rolon, 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*. San Diego, USA, 2008, pp. 129–132. doi: [10.1109/ICIP.2008.4711708](https://doi.org/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](https://doi.org/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. 2020.
- [TH2] Xavier Alameda-Pineda. "Egocentric Audio-Visual Scene Analysis, A Machine Learning and Signal Processing Approach". PhD Thesis. Université Joseph Fourier, 2013.
- [TH3] Xavier Alameda-Pineda. "Finding audio-visual events with a robot head". MSc Thesis. Université Joseph Fourier/ENSIMAG/Grenoble INP, 2010.
- [TH4] Xavier Alameda-Pineda. "Entropy Coding for Image Compression Based on Generalized Lifting and SPECK". MSc Thesis (equivalent). UPC BarcelonaT-ech, 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](https://doi.org/10.1016/C2017-0-01387-3).

#### PATENTS

- [P1] Nicu Sebe, Xavier Alameda-Pineda, Sergey Tulyakov, Elisa Ricci, Lijun Yin, and Jeffrey F. Cohn. "Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions". Patent Application US 15631346. 2017.

#### 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. Chap. 1, pp. 1–10. doi: [10.1016/B978-0-12-814601-9.00011-0](https://doi.org/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](https://doi.org/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](https://doi.org/10.1016/B978-0-12-809276-7.00017-5).