

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

Research Scientist at the **Perception Team**

Artificial Intelligence co-Chair at **MIAI**

on *Audio-visual machine perception and interaction for companion robots*

**INRIA Grenoble Rhône-Alpes, Université Grenoble Alpes**

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



## ACADEMIC AND PROFESSIONAL QUALIFICATIONS

2020 — **ELLIS Membmer** (**ELLIS**, USA)  
2019 — **IEEE Senior Membmer** (IEEE, USA)  
2014 — **Qualification for Maître de Conférences Section CNU 61** (**MENESR**, Paris, France)  
2013 — **PhD in Mathematics and Computer Science** (Perception Team, INRIA, **UJF**, Grenoble, France)  
2010 — **Masters in Computer Science (Graphics, Vision and Robotics)** (**MoSIG**, UFR **IM2AG**, UJF, **Grenoble-INP**, Grenoble, France)  
2009 — **Masters (equivalent) in Telecommunications Engineering** (**ETSETB - BarcelonaTECH**, Barcelona, Spain)  
2008 — **Masters (equivalent) in Mathematics** (**FME - BarcelonaTECH**, Barcelona, Spain)

## RESEARCH EXPERIENCE

**Research Scientist** [REF.: DR. RADU HORAUD] (Perception Team, INRIA, 2016-Now, Grenoble, France)  
**Postdoctoral Fellow** [REF.: PROF. NICU SEBE] (**MHUG, DISI**, University of Trento, 2014-2016, Trento, Italy)  
**Postdoctoral Fellow** [REF.: PROFS. LAURENT GIRIN/RADU HORAUD] (MAGIC/Perception Teams, GIPSA-LAB/INRIA, 2013-2014, Grenoble, France)  
**PhD Thesis: Egocentric Audio-Visual Scene Analysis, A Machine Learning and Signal Processing Approach** [REF.: DR. RADU HORAUD] (Perception Team, INRIA, UJF, 2010-2013, Grenoble, France)  
**Masters Thesis: Finding audio-visual event with a robot head** [REF.: DR. RADU HORAUD] (MoSIG, UFR IM2AG, UJF and Perception Team, INRIA, 2009-10, Grenoble, France)  
**Research collaboration grant** [REF.: PROF. PHILIPPE SALEMBIER] (**IPG** in the **STC** at UPC, Summer of 2007 and of 2008, Barcelona, Spain)  
**Research collaboration grant for young researchers** [REF.: PROF. XAVIER CABRÉ] (**MEC** and **MA1** at UPC, 2007-08, Barcelona, Spain)

## AWARDS AND HONORS

**ACM SIGMM Rising Star Award 2018** for *Contributions to Multimodal Social Behavior Understanding*  
at **SIGMM Web Site**  
**Novel Technology Award Finalist** for *Tracking a Varying Number of People with a Visually-Controlled Robotic Head* [**C16**]  
at **International Conference on Intelligent Robots and Systems'17**  
**Best scientific paper award** for *Multi-Paced Dictionary Learning for Cross-Domain Retrieval and Recognition* [**C29**]  
at **IAPR International Conference on Pattern Recognition'16**  
**Best paper award** for *Analysing Free-standing Conversational Groups: A Multimodal Approach* [**C30**]  
at **ACM International Conference on Multimedia'15**  
**Best student paper award** for *A Variational EM Algorithm for the Separation of Moving Sound Sources* [**C31**]  
at **IEEE Workshop on Applications of Signal Processing to Audio and Acoustics'15**  
**Outstanding paper award** for *Finding Audio-visual Events in Informal Social Gatherings* [**C41**]  
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

### PhD Supervision:

Name	Period	Affiliation	(Co-)Supervisor
<b>PhD Supervisor – Directeur de thèse</b>			
<b>Xiaoyu Lin</b>	2020-	Perception Team, INRIA, Grenoble, France	Prof. Laurent Girin
<b>Gaétan Lepage</b>	2020-	Perception Team, INRIA, Grenoble, France	Prof. Laurent Girin
<b>Wen Guo</b>	2019-	Perception Team, INRIA, Grenoble, France	Dr. Francesc Moreno-Noguer
<b>PhD Co-supervisor – Co-encadrant de thèse</b>			
<b>Anand Ballou</b>	2019-	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Louis Airale</b>	2019-	Perception Team, INRIA, Grenoble, France	Prof. Dominique Vaufreydaz
<b>Xiaoyu Bie</b>	2019-	Perception Team, INRIA, Grenoble, France	Prof. Laurent Girin
<b>Yihong Xu</b>	2018-	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Guillaume Delorme</b>	2017-	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Aliaksandr Siahorin</b>	2016-17	MHUG, University of Trento, Trento, Italy	Prof. Nicu Sebe
<b>Wei Wang</b>	2016-18	MHUG, University of Trento, Trento, Italy	Prof. Nicu Sebe
<b>Andrea Pilzer</b>	2016-17	MHUG, University of Trento, Trento, Italy	Prof. Nicu Sebe
<b>Dan Xu</b>	2014-17	MHUG, University of Trento, Trento, Italy	Prof. Nicu Sebe
<b>Sergey Tulyakov</b>	2015-17	MHUG, University of Trento, Trento, Italy	Prof. Nicu Sebe
<b>Occasional collaboration – Suivi de thèse</b>			
<b>Stéphane Lathuilière</b>	2016-17	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Vicent Drouard</b>	2016-17	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Yutong Ban</b>	2016-19	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Dionyssos Kounades-Bastian</b>	2014-17	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Israel-Dejene Gebru</b>	2014-17	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Maxime Janvier</b>	2013-14	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud

### Master Thesis Supervision:

Name	Period	Affiliation	(Co-)Supervisor
<b>Viet Nhat Nguyen</b>	2020	Perception Team, INRIA, Grenoble, France	Dr. Mostafa Sadeghi
<b>Álvaro González Jiménez</b>	2020	Perception Team, INRIA, Grenoble, France	Prof. Stéphane Lathuilière
<b>Predrag Pilipovic</b>	2020	Perception Team, INRIA, Grenoble, France	
<b>Vadim Sushko</b>	2019	Perception Team, INRIA, Grenoble, France	
<b>Elsa Marie</b>	2019	Perception Team, INRIA, Grenoble, France	
<b>Divya Grover</b>	2018	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Aliaksandr Siarohin</b>	2016-17	MHUG, University of Trento, Trento, Italy	Prof. Nicu Sebe
<b>Alessio Xompero</b>	2014	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Israel-Dejene Gebru</b>	2013	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud
<b>Maxime Janvier</b>	2011-12	Perception Team, INRIA, Grenoble, France	Dr. Radu Horaud

## TEACHING

**Category Learning and Object Recognition (9h)** [REF.: DR. KARTEEK ALAHARI] (ENSIMAG/UGA, 2019-2020, Grenoble, France)  
**Advanced Learning Models (9h)** [REF.: DR. JULIEN MAIRAL] (ENSIMAG/UGA, 2019-2020, Grenoble, France)  
**Fundamentals of Probabilistic Data Mining (13h)** [REF.: DR. THOMAS HUEBER] (ENSIMAG/UGA, 2019-2020, Grenoble, France)  
**Fundamentals of Probabilistic Data Mining (5h)** [REF.: PROF. JEAN-BAPTISTE DURAND] (ENSIMAG/UGA, 2018-2019, Grenoble, France)  
**Algebra and Geometry (128h)** [REF.: PROFS. FRANÇOIS DAHMANI AND STÉPHANIE POUGET] (**DLST**, UJF, 2012-13, Grenoble, France)

## TUTORIALS & INVITED TALKS

**Audio-visual variational speech enhancement (Sep'20)**  
at **Intelligent Sensing Summer School**  
**Choosing wisely your deep training loss – seminar (March'20)**  
at **Universidade NOVA de Lisboa**  
**Artificial Intelligence for Social Robots in Gerontological Healthcare (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 – Seminar** (*Jul'19*)  
at University of Trento

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

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

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

**Audio-Visual Multiple Speaker with Robotic Platforms** (*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 – seminar** (*Jun'16*)  
at [Télécom-ParisTech](#)

**Variational EM and non-linear optimization for multi-sensor scene analysis – seminar** (*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** (*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](#)

## COMMUNITY SERVICE

**Associated Editor:** ACM TOMM (2019-)

**Special issue organisation:** ICJV'19, ACM TOMM'18

**Chairing:** ACM MM'22 program chair, IEEE WACV'21 area chair, IAPR ICPR'20 area chair, ACM MM'19 & '20 area chair, ICIAP'19 area chair, ICCV'17 area chair

**Workshop organisation:** ACM MM'17, '18, '19 & '20, ICCV'19, CVPR'18 & 19, ECCV'18, MediaEval'19

**Special session organisation:** IEEE ICIP'19, ACM ICMR'17

**Reviewing for International Journals:** IEEE TPAMI, IEEE TIP, IEEE TSP, IEEE/ACM TASLP, IJRR, CVIU.

**Reviewing for International Conferences:** IEEE CVPR'17-19, NIPS'16-19, ICLR'18-19, ECCV'16,18, ACM MM'16-19, ICMR'12-16.

**PhD Reviewing:** Wei Wang (University of Trento'18), Dan Xu (University of Trento'18), Daniek Brink (University of Stellenbosh'16), José Velasco (University of Alcalá'16).

**PhD Committee:** Yutong Ban (University Grenoble-Alpes'19), Irtiza Hasan (University of Verona'19), Theodoros Tsesmelis (University of Verona'19), Israel-Dejene Gebru (University of Grenoble-Alpes'18), Dionyssos Kounades-Bastian (University of Grenoble Alpes'17).

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

## ACADEMIC COLLABORATIONS

**IRI-UPC:** multi-person full-body 3D pose estimation.

**Fondazione Bruno Kessler:** analysis of free-standing conversational groups, head/body pose estimation.

**TUM:** multi-person tracking.

**Bar Ilan University:** sound source separation with variational probabilistic models.

**University of Trento:** human behavior analysis, low-rank methods, recognition of subjective properties.

**GIPSA-LAB:** probabilistic models for acoustico-articulatory inversion adaptation.

## SOFTWARE AND DATASETS

The [SALSA](#) dataset, presented in [J11], used in [C30].

The [Geometric Time Delay Estimation](#) MATLAB toolbox, developed in [C37, C34, J17].

The [APP IDDN](#) for “Association of Audio Cues with 3D locations”, library developed in [C41, C40, J15].

The [RAVEL](#) dataset, presented in [J18], used in [C41].

## NATURAL & PROGRAMMING LANGUAGES

**Catalan:** Mother tongue. **Spanish:** Expert. **English & French:** Professional/Teaching. **Italian:** Intermediate.

**C:** Mother tongue. **Matlab & LaTeX:** Expert. **C++ & UNIX Shell:** Fluent. **Java & Python:** Beginner.

## PUBLICATIONS

### BOOKS

- [B1] Xavier Alameda-Pineda, Elisa Ricci, and Nicu Sebe. *Multimodal Behavior Analysis in the Wild: Advances and Challenges*. Elsevier, 2018.

### 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.
- [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.
- [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.

### JOURNAL ARTICLES

- [J1] 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* (2020). doi: [10.1109/TPAMI.2019.2953020](https://doi.org/10.1109/TPAMI.2019.2953020).
- [J2] 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).
- [J3] 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* (2020). doi: [10.1109/TMM.2019.2963621](https://doi.org/10.1109/TMM.2019.2963621).
- [J4] 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).
- [J5] 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).
- [J6] 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).
- [J7] 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).
- [J8] 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).
- [J9] 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).
- [J10] 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).
- [J11] 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).
- [J12] 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). URL: <http://arxiv.org/abs/1509.01520>.
- [J13] 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). URL: <http://arxiv.org/abs/1509.01509>.
- [J14] 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). URL: <http://arxiv.org/abs/1510.04595>.
- [J15] 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).
- [J16] 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).
- [J17] 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).

- [J18] 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] 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.
- [C2] 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.
- [C3] 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.
- [C4] 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.
- [C5] 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.
- [C6] 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: [3343031.3350590](https://doi.org/10.1145/3343031.3350590).
- [C7] 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 MM*. Nice, France, 2019.
- [C8] 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.
- [C9] 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.
- [C10] 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.
- [C11] 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.
- [C12] 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.
- [C13] 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.
- [C14] Xavier Alameda-Pineda, Andrea Pilzer, Dan Xu, Nicu Sebe, and Elisa Ricci. "Viraliency: Pooling Local Virality". In: *IEEE International Conference on Computer Vision and Pattern Recognition*. Hawaii, USA, 2017. doi: [10.1109/CVPR.2017.59](https://doi.org/10.1109/CVPR.2017.59).
- [C15] 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).
- [C16] 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.
- [C17] 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.
- [C18] 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.
- [C19] 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.
- [C20] 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.
- [C21] 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.
- [C22] 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.
- [C23] 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.
- [C24] 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.
- [C25] 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.
- [C26] 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.



- [C27] 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.
- [C28] 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.
- [C29] 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.
- [C30] 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.
- [C31] 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.
- [C32] 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.
- [C33] 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.
- [C34] 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.
- [C35] 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.
- [C36] 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.
- [C37] Xavier Alameda-Pineda and Radu Horaud. "Geometrically-constrained Robust Time Delay Estimation Using Non-coplanar Microphone Arrays". In: *European Signal Processing Conference*. Bucharest, Romania, 2012, pp. 1309–1313.
- [C38] 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.
- [C39] 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.
- [C40] 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.
- [C41] 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.
- [C42] 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.

## PUBLIC SOFTWARE REPOSITORIES

- [SR1] Xavier Alameda-Pineda, Andrea Pilzer, Dan Xu, Nicu Sebe, and Elisa Ricci. *Learned Average Pooling*. [https://github.com/xavirema/lena\\_pooling](https://github.com/xavirema/lena_pooling). 2017.
- [SR2] Thomas Hueber, Xavier Alameda-Pineda, and Laurent Girin. *The Integrated Cascaded Gaussian Mixture Regressor*. <https://github.com/thueber/cgmr.git>. 2017.
- [SR3] Xavier Alameda-Pineda, Elisa Ricci, Yan Yan, and Nicu Sebe. *Non-Linear Matrix Completion*. <https://github.com/xavirema/nlmc>. 2016.
- [SR4] Israel-Dejene Gebru, Xavier Alameda-Pineda, Florence Forbes, and Radu Horaud. *EM for weighted-data GMM*. <http://perception.inrialpes.fr/people/Gebru/code/WD-EM.zip>. 2016.
- [SR5] Xavier Alameda-Pineda and Radu Horaud. *The gtde MATLAB toolbox*. <https://team.inria.fr/perception/the-gtde-matlab-toolbox/>. 2014.
- [SR6] Xavier Alameda-Pineda, Antoine Deleforge, Soraya Arias, and Radu Horaud. *Association of Audio Cues with 3D locations library version 1.0*. IDDN FR 001 420023 000 S P 2014 000 31235. Apr. 2013.
- [SR7] Xavier Alameda-Pineda, Antoine Deleforge, Jordi Sanchez-Riera, Johannes Wienke, Soraya Arias, Sebastian Wrede, and Radu Horaud. *Auditory, visual and audio-visual speaker localization with the NAO robot*. <http://code.humavips.eu>. 2013.

## PREPRINTS

- [U1] Xavier Alameda-Pineda, Vincent Drouard, and Radu Horaud. "Variational Inference and Learning of Piecewise-linear Dynamical Systems". Submitted to IEEE Transactions on Neural Networks and Learning Systems. 2020.
- [U2] Guillaume Delorme, Yutong Ban, Guillaume Sarrazin, and Xavier Alameda-Pineda. "ODA-Track: Online Deep Appearance for Robotic Multiple Person Tracking". 2020.
- [U3] Laurent Girin, Simon Leglaive, Xiaoyu Bie, Julien Diard, Thomas Hueber, and Xavier Alameda-Pineda. "Dynamical Variational Autoencoders: A Comprehensive Review". 2020.
- [U4] Yahui Liu, Marco De Nadai, Jian Yao, Nicu Sebe, Bruno Lepri, and Xavier Alameda-Pineda. "GMM-UNIT: Unsupervised Multi-Domain and Multi-Modal Image-to-Image Translation via Attribute Gaussian Mixture Modeling". 2020.

- [U5] Viet-Nhat Nguyen, Mostafa Sadeghi, Elisa Ricci, and Xavier Alameda-Pineda. "Deep Variational Generative Models for Audio-visual Speech Separation". 2020.
- [U6] Mostafa Sadeghi and Xavier Alameda-Pineda. "Mixture of Inference Networks for VAE-based Audio-visual Speech Enhancement". Submitted to IEEE TSP. 2020.