Xavier Alameda-Pineda, Eng, PhD

Research Scientist at the Perception Team[1]

INRIA Grenoble Rhône-Alpes[2], Université Grenoble Alpes[3]

xavirema@xavirema.eu

Diplomas

2014 – **Qualification for *Maître de Conférences* Section CNU 61** (MENESR[4], Paris, France)
2013 – **PhD in Mathematics and Computer Science** (Perception Team, INRIA, UJF[5], Grenoble, France)
2010 – **Masters in Computer Science (Graphics, Vision and Robotics)** (MoSIG[6], UFR IM2AG[7], UJF, Grenoble-INP[8] , Grenoble, France)
2009 – **Masters (equivalent) in Telecomunications Engineering** (ETSETB[9] - UPC[10], Barcelona, Spain)
2008 – **Masters (equivalent) in Mathematics** (FME[11] - UPC, Barcelona, Spain)

Research Experience

Research Scientist [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2016-Now, Grenoble, France)
Postdoctoral Fellow [Ref.: Prof. Nicu Sebe] (MHUG[12], DISI[13], 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[14] in the STC[15] at UPC, Summer of 2007 and of 2008, Barcelona, Spain)
Research collaboration grant for young researchers [Ref.: Prof. Xavier Cabré] (MEC[16] and MA1[17] at UPC, 2007-08, Barcelona, Spain)

PhD Co-advisor:

Stephane Lathuillere [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2016-17, Grenoble, France)
Yutong Ban [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2016-17, Grenoble, France)
Andrea Pilzer [Ref.: Prof. Nicu Sebe] (MHUG, University of Trento, 2016-17, Trento, Italy)
Dan Xu [Ref.: Prof. Nicu Sebe] (MHUG, University of Trento, 2015-17, Trento, Italy)
Sergey Tulyakov [Ref.: Prof. Nicu Sebe] (MHUG, University of Trento, 2015-17, Trento, Italy)
Dionyssos Kounades-Bastian [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2014-17, Grenoble, France)
Israel-Dejene Gebru [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2014-17, Grenoble, France)
Maxime Janvier [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2013-14, Grenoble, France)

Master Thesis Co-advisor:

Guillaume Delorme [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2017, Grenoble, France)
Aliaksandr Siarohin [Ref.: Prof. Nicu Sebe] (MHUG, Univerity of Trento, 2016, Trento, Italy)
Alessio Xompero [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2014, Grenoble, France)
Israel-Dejene Gebru [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2013, Grenoble, France)
Maxime Janvier [Ref.: Dr. Radu Horaud] (Perception Team, INRIA, 2011-12, Grenoble, France)

Teaching Algebra and Geometry (128h) [Ref.: Profs. François Dahmani and Stéphanie Pouget] (DLST[18], UJF, 2012-13, Grenoble, France)

Awards and Honors

Best scientific paper award for *Multi-Paced Dictionary Learning for Cross-Domain Retrieval and Recognition* [C14]
at IAPR International Conference on Pattern Recognition'16

Best paper award for *Analysing Free-standing Conversational Groups: A Multimodal Approach* [C15]
at ACM International Conference on Multimedia'15

Best student paper award for *A Variational EM Algorithm for the Separation of Moving Sound Sources* [C16]
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* [C26]
at IEEE/ACM International Conference on Multimodal Interaction'11

Invited Talks

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 – Multimodal Rising Star Lecture (*Oct'15*)
at SIGMM Inaugural Workshop on Multimedia Frontiers

Community service

Chairing: IEEE ICCV'17 area chair, ACM MM'17 MUSA2 Workshop Chair, ACM ICMR'17 special session chair.
Revieweing for International Journals: IEEE TPAMI, IEEE TIP, IEEE TSP, IEEE/ACM TASLP, IJRR, CVIU.
Revieweing for International Conferences: IEEE CVPR'17, NIPS'16, ECCV'16, ACM MM'16, ICMI'12-16.
PhD Reviewing: Daniek Brink (University of Stellenbosh'16), José Velasco (University of Alcalá'16).
PhD Committee: Dionyssos Kounades-Bastian (University of Grenoble Alpes'17).

Academic Collaborations

GIPSA-LAB: probabilistic models for acoustico-articulatory inversion adaptation.
Fondazione Bruno Kessler: analysis of free-standing conversational groups, head/body pose estimation.
Bar Ilan University: sound source separation with variational probabilistic models.
University of Trento: human behavior analysis, low-rank methods, recognition of subjective properties .

Project Participation

VHIA ERC Advanced Grant: multi-modal scene analysis, audio-visual tracking, pose estimation.
ACANTO Horizon2020 project: detection and tracking of speakers, facial expression recognition.
HUMAVIPS EU-FP7 project: audio-visual speaker detection and localization, final demonstrator.

Software and Datasets

The SALSA dataset, presented in [J2], used in [C15].
The Geometric Time Delay Estimation MATLAB toolbox, developed in [C22, C19, J8].
The APP[19] IDDN[20] for "Association of Audio Cues with 3D locations", library developed in [C26, C25, J6].
The RAVEL dataset, presented in [J9], used in [C26].

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

Journal Articles

[J1] 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.

[J2] **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.

[J3] 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. URL: http://arxiv.org/abs/1509.01520.

[J4] 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. URL: http://arxiv.org/abs/1509.01509.

[J5] 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. URL: http://arxiv.org/abs/1510.04595.

[J6] **Xavier Alameda-Pineda** and Radu Horaud. "Vision-Guided Robot Hearing". In: *International Journal of Robotics Research* 34.4-5 (2015), pp. 437–456.

[J7] 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.

[J8] **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.

[J9] **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.

Book Chapters

[CH1] **Xavier Alameda-Pineda**, Elisa Ricci, and Nicu Sebe. "Multimodal analysis of free-standing conversational groups". In: *Research Frontiers of Multimedia*. Ed. by Shih-Fu Chang. Morgan and Claypool, 2016.

[CH2] **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.

Conference Papers

[C1] **Xavier Alameda-Pineda**, Andrea Pilzer, Dan Xu, Nicu Sebe, and Elisa Ricci. "Viraliency: Pooling Local Viraliry". In: *IEEE International Conference on Computer Vision and Pattern Recognition*. Hawai, USA, 2017.

[C2] **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.

[C3] 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.

[C4] 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.

[C5] 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.

[C6] 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.

[C7] 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.

[C8] **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.

[C9] 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.

[C10] 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.

[C11] 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.

[C12] 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.

[C13] 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.

[C14] 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.

[C15] **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.

[C16] 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.

[C17] 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.

[C18] 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.

[C19] **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.

[C20] **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.

[C21] 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.

[C22] **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.

[C23] 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.

[C24] 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.

[C25] 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.

[C26] **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.

[C27] 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.

Notes

[1] Perception Team's web site

[2] Institut National de Recherche en Informatique et Automatique, Siège de Grenoble Rhône-Alpes

[3] Université Grenoble Alpes

[4] Ministère de l'Éducation National, l'Enseignement Supérieur et la Recherche

[5] Université Joseph Fourier, Grenoble 1

[6] Master of Science in Informatics at Grenoble

[7] Unité de Formation et Recherche en Informatique, Mathématiques et Mathématiques Appliquées de Grenoble

[8] Grenoble-INP

[9] Escola Tècnica Superior d'Enginyeria de Telecomunicacions de Barcelona

[10] Universitat Politècnica de Catalunya – BarcelonaTECH

[11] Facultat de Matemàtiques i Estadística

[12] Multimodal Human Understanding Group

[13] Department of Information Engineering and Computer Science

[14] Image Processing Group

[15] Signal Theory and Communications Department

[16] Ministerio de Educación y Ciencia

[17] Department of Applied Mathematics 1

[18] Département License Sciences et Technologies

[19] Agence pour la Protection des Programmes

[20] Inter Deposit Digital Number