




# Xavier Alameda-Pineda, Eng, PhD

Research Scientist at the Perception Team<sup>1</sup>

INRIA Grenoble Rhône-Alpes<sup>2</sup>, Université Grenoble Alpes<sup>3</sup>

✉ [xavirema@xavirema.eu](mailto:xavirema@xavirema.eu)   
 



## ACADEMIC DIPLOMAS

- 2014 – **Qualification for *Maître de Conférences* Section CNU 61** (MENESR<sup>4</sup>, Paris, France)
- 2013 – **PhD in Mathematics and Computer Science** (Perception Team, INRIA, UJF<sup>5</sup>, Grenoble, France)
- 2010 – **Masters in Computer Science (Graphics, Vision and Robotics)** (MoSIG<sup>6</sup>, UFR IM2AG<sup>7</sup>, UJF, Grenoble-INP<sup>8</sup>, Grenoble, France)
- 2009 – **Masters (equivalent) in Telecommunications Engineering** (ETSETB<sup>9</sup> - UPC<sup>10</sup>, Barcelona, Spain)
- 2008 – **Masters (equivalent) in Mathematics** (FME<sup>11</sup> - 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<sup>12</sup>, DISI<sup>13</sup>, 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<sup>14</sup> in the STC<sup>15</sup> at UPC, Summer of 2007 and of 2008, Barcelona, Spain)
- Research collaboration grant for young researchers** [REF.: PROF. XAVIER CABRÉ] (MEC<sup>16</sup> and MA1<sup>17</sup> at UPC, 2007-08, Barcelona, Spain)

## TEACHING AND SUPERVISION

### PhD Co-advisor:

- Stéphane Lathuilière** [REF.: DR. RADU HORAUD] (Perception Team, INRIA, 2016-17, Grenoble, France)
- Vicent Drouard** [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:

- Divya Grover** [REF.: DR. RADU HORAUD] (Perception Team, INRIA, 2017, Grenoble, France)
- Aliaksandr Siarohin** [REF.: PROF. NICU SEBE] (MHUG, University of Trento, 2016-17, 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<sup>18</sup>, UJF, 2012-13, Grenoble, France)

## AWARDS AND HONORS

- Best scientific paper award** for *Multi-Paced Dictionary Learning for Cross-Domain Retrieval and Recognition* [C11] at [IAPR International Conference on Pattern Recognition'16](#)
- Best paper award** for *Analysing Free-standing Conversational Groups: A Multimodal Approach* [C12] at [ACM International Conference on Multimedia'15](#)
- Best student paper award** for *A Variational EM Algorithm for the Separation of Moving Sound Sources* [C13] 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* [C23] at [IEEE/ACM International Conference on Multimodal Interaction'11](#)

## TUTORIALS & 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 – Workshop on Multimedia Frontiers Lecture** (*Oct'15*)

at [SIGMM Inaugural Workshop on Multimedia Frontiers](#)

## COMMUNITY SERVICE

Chairing: IEEE ICCV'17 area chair, ACM ICMR'17 special session organizer.

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

Reviewing for International Conferences: IEEE CVPR'17, NIPS'16, ECCV'16, ACM MM'16, ICMJ'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 [C12].

The **Geometric Time Delay Estimation** MATLAB toolbox, developed in [C19, C16, J8].

The APP<sup>19</sup> IDDN<sup>20</sup> for “Association of Audio Cues with 3D locations”, library developed in [C23, C22, J6].

The **RAVEL** dataset, presented in [J9], used in [C23].

## NATURAL & PROGRAMMING LANGUAGES

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

**C**: Mother tongue. **Matlab & L<sup>A</sup>T<sub>E</sub>X**: Expert. **C++ & UNIX Shell**: Fluent. **Java & Python**: Beginner.

## PUBLICATIONS

### BOOKS

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

### 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.

### 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](https://doi.org/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] Siley 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](https://doi.org/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 Convolutional 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.

### CONFERENCE PAPERS

- [C1] 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.
- [C2] 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.
- [C3] Dionyssos Kounades-Bastian, Laurent Girin, Xavier Alameda-Pineda, Sharon Gannot, and Radu Horaud. “An EM algorithm for joint source separation and diarisation of multichannel convolutional mixtures”. In: *IEEE International Conference on Audio, Speech and Signal Processing*. New Orleans, USA, 2017.
- [C4] 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.
- [C5] 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.
- [C6] Yutong Ban, Siley 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.
- [C7] 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.
- [C8] 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.
- [C9] 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.
- [C10] 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.

- [C11] 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.
- [C12] 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.
- [C13] 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.
- [C14] 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.
- [C15] 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.
- [C16] 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.
- [C17] 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.
- [C18] 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.
- [C19] 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.
- [C20] 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.
- [C21] 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.
- [C22] 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.
- [C23] 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.
- [C24] 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

<sup>1</sup>Perception Team’s web site

<sup>2</sup>Institut National de Recherche en Informatique et Automatique, Siège de Grenoble Rhône-Alpes

<sup>3</sup>Université Grenoble Alpes

<sup>4</sup>Ministère de l’Éducation National, l’Enseignement Supérieur et la Recherche

<sup>5</sup>Université Joseph Fourier, Grenoble 1

<sup>6</sup>Master of Science in Informatics at Grenoble

<sup>7</sup>Unité de Formation et Recherche en Informatique, Mathématiques et Mathématiques Appliquées de Grenoble

<sup>8</sup>Grenoble-INP

<sup>9</sup>Escola Tècnica Superior d’Enginyeria de Telecomunicacions de Barcelona

<sup>10</sup>Universitat Politècnica de Catalunya – BarcelonaTECH

<sup>11</sup>Facultat de Matemàtiques i Estadística

<sup>12</sup>Multimodal Human Understanding Group

<sup>13</sup>Department of Information Engineering and Computer Science

<sup>14</sup>Image Processing Group

<sup>15</sup>Signal Theory and Communications Department

<sup>16</sup>Ministerio de Educación y Ciencia

<sup>17</sup>Department of Applied Mathematics 1

<sup>18</sup>Département License Sciences et Technologies

<sup>19</sup>Agence pour la Protection des Programmes

<sup>20</sup>Inter Deposit Digital Number