Xavier Alameda-Pineda

List of Publications by Year
in descending order

33 papers	**880** citations
33 all docs	**33** docs citations

687363

13 h-index

794594

19 g-index

33 times ranked

876 citing authors

#	Article	IF	Citations
1	A Comprehensive Analysis of Deep Regression. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42, 2065-2081.	13.9	148
2	SALSA: A Novel Dataset for Multimodal Group Behavior Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38, 1707-1720.	13.9	107
3	EM Algorithms for Weighted-Data Clustering with Application to Audio-Visual Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38, 2402-2415.	13.9	86
4	A Geometric Approach to Sound Source Localization from Time-Delay Estimates. IEEE/ACM Transactions on Audio Speech and Language Processing, 2014, 22, 1082-1095.	5.8	64
5	Analyzing Free-standing Conversational Groups. , 2015, , .		61
6	Audio-Visual Speech Enhancement Using Conditional Variational Auto-Encoders. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 1788-1800.	5.8	39
7	Automatic animation of an articulatory tongue model from ultrasound images of the vocal tract. Speech Communication, 2017, 93, 63-75.	2.8	38
8	An on-line variational Bayesian model for multi-person tracking from cluttered scenes. Computer Vision and Image Understanding, 2016, 153, 64-76.	4.7	33
9	Online Localization and Tracking of Multiple Moving Speakers in Reverberant Environments. IEEE Journal on Selected Topics in Signal Processing, 2019, 13, 88-103.	10.8	29
10	Vision-guided robot hearing. International Journal of Robotics Research, 2015, 34, 437-456.	8.5	28
11	RAVEL: an annotated corpus for training robots with audiovisual abilities. Journal on Multimodal User Interfaces, 2013, 7, 79-91.	2.9	27
12	Variational Bayesian Inference for Audio-Visual Tracking of Multiple Speakers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43, 1761-1776.	13.9	27
13	Active-speaker detection and localization with microphones and cameras embedded into a robotic head. , 2013, , .		20
14	Speaker-Adaptive Acoustic-Articulatory Inversion Using Cascaded Gaussian Mixture Regression. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, 23, 2246-2259.	5.8	18
15	Viraliency: Pooling Local Virality. , 2017, , .		16
16	Tracking Multiple Audio Sources With the von Mises Distribution and Variational EM. IEEE Signal Processing Letters, 2019, 26, 798-802.	3.6	16
17	Exploiting the Complementarity of Audio and Visual Data in Multi-speaker Tracking. , 2017, , .		15
18	Cross-Paced Representation Learning With Partial Curricula for Sketch-Based Image Retrieval. IEEE Transactions on Image Processing, 2018, 27, 4410-4421.	9.8	14

#	Article	IF	Citations
19	Mixture of Inference Networks for VAE-Based Audio-Visual Speech Enhancement. IEEE Transactions on Signal Processing, 2021, 69, 1899-1909.	5.3	13
20	Learning How to Smile: Expression Video Generation With Conditional Adversarial Recurrent Nets. IEEE Transactions on Multimedia, 2020, 22, 2808-2819.	7.2	12
21	Academic Coupled Dictionary Learning for Sketch-based Image Retrieval. , 2016, , .		11
22	Tracking a varying number of people with a visually-controlled robotic head. , 2017, , .		11
23	Robust Unsupervised Audio-Visual Speech Enhancement Using a Mixture of Variational Autoencoders. , 2020, , .		10
24	DeepGUM: Learning Deep Robust Regression with a Gaussian-Uniform Mixture Model. Lecture Notes in Computer Science, 2018, , 205-221.	1.3	10
25	Accounting for Room Acoustics in Audio-Visual Multi-Speaker Tracking. , 2018, , .		9
26	Extending the Cascaded Gaussian Mixture Regression Framework for Cross-Speaker Acoustic-Articulatory Mapping. IEEE/ACM Transactions on Audio Speech and Language Processing, 2017, 25, 662-673.	5.8	4
27	Switching Variational Auto-Encoders for Noise-Agnostic Audio-Visual Speech Enhancement. , 2021, , .		4
28	Audio-Visual Variational Fusion for Multi-Person Tracking with Robots. , 2019, , .		4
29	Variational Inference and Learning of Piecewise Linear Dynamical Systems. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33, 3753-3764.	11.3	3
30	SALSA: A Multimodal Dataset for the Automated Analysis of Free-Standing Social Interactions. , 2017, , 321-340.		1
31	Deep Variational Generative Models for Audio-Visual Speech Separation. , 2021, , .		1
32	The Impact of Removing Head Movements on Audio-Visual Speech Enhancement. , 2022, , .		1
33	ODANet: Online Deep Appearance Network for Identity-Consistent Multi-person Tracking. Lecture Notes in Computer Science, 2021, , 803-818.	1.3	0