

Xavier Alameda-Pineda

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/4959239/publications.pdf>

Version: 2024-02-01

33
papers

880
citations

687363

13
h-index

794594

19
g-index

33
all docs

33
docs citations

33
times ranked

876
citing authors

#	ARTICLE	IF	CITATIONS
1	Variational Inference and Learning of Piecewise Linear Dynamical Systems. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33, 3753-3764.	11.3	3
2	The Impact of Removing Head Movements on Audio-Visual Speech Enhancement. , 2022, , .		1
3	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
4	Mixture of Inference Networks for VAE-Based Audio-Visual Speech Enhancement. IEEE Transactions on Signal Processing, 2021, 69, 1899-1909.	5.3	13
5	Switching Variational Auto-Encoders for Noise-Agnostic Audio-Visual Speech Enhancement. , 2021, , .		4
6	ODANet: Online Deep Appearance Network for Identity-Consistent Multi-person Tracking. Lecture Notes in Computer Science, 2021, , 803-818.	1.3	0
7	Deep Variational Generative Models for Audio-Visual Speech Separation. , 2021, , .		1
8	A Comprehensive Analysis of Deep Regression. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42, 2065-2081.	13.9	148
9	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
10	Learning How to Smile: Expression Video Generation With Conditional Adversarial Recurrent Nets. IEEE Transactions on Multimedia, 2020, 22, 2808-2819.	7.2	12
11	Robust Unsupervised Audio-Visual Speech Enhancement Using a Mixture of Variational Autoencoders. , 2020, , .		10
12	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
13	Tracking Multiple Audio Sources With the von Mises Distribution and Variational EM. IEEE Signal Processing Letters, 2019, 26, 798-802.	3.6	16
14	Audio-Visual Variational Fusion for Multi-Person Tracking with Robots. , 2019, , .		4
15	Accounting for Room Acoustics in Audio-Visual Multi-Speaker Tracking. , 2018, , .		9
16	Cross-Paced Representation Learning With Partial Curricula for Sketch-Based Image Retrieval. IEEE Transactions on Image Processing, 2018, 27, 4410-4421.	9.8	14
17	DeepGUM: Learning Deep Robust Regression with a Gaussian-Uniform Mixture Model. Lecture Notes in Computer Science, 2018, , 205-221.	1.3	10
18	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

#	ARTICLE	IF	CITATIONS
19	Automatic animation of an articulatory tongue model from ultrasound images of the vocal tract. <i>Speech Communication</i> , 2017, 93, 63-75.	2.8	38
20	Tracking a varying number of people with a visually-controlled robotic head. , 2017, , .		11
21	Viraliency: Pooling Local Virality. , 2017, , .		16
22	Exploiting the Complementarity of Audio and Visual Data in Multi-speaker Tracking. , 2017, , .		15
23	SALSA: A Multimodal Dataset for the Automated Analysis of Free-Standing Social Interactions. , 2017, , 321-340.		1
24	EM Algorithms for Weighted-Data Clustering with Application to Audio-Visual Scene Analysis. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2016, 38, 2402-2415.	13.9	86
25	An on-line variational Bayesian model for multi-person tracking from cluttered scenes. <i>Computer Vision and Image Understanding</i> , 2016, 153, 64-76.	4.7	33
26	Academic Coupled Dictionary Learning for Sketch-based Image Retrieval. , 2016, , .		11
27	SALSA: A Novel Dataset for Multimodal Group Behavior Analysis. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i> , 2016, 38, 1707-1720.	13.9	107
28	Vision-guided robot hearing. <i>International Journal of Robotics Research</i> , 2015, 34, 437-456.	8.5	28
29	Analyzing Free-standing Conversational Groups. , 2015, , .		61
30	Speaker-Adaptive Acoustic-Articulatory Inversion Using Cascaded Gaussian Mixture Regression. <i>IEEE/ACM Transactions on Audio Speech and Language Processing</i> , 2015, 23, 2246-2259.	5.8	18
31	A Geometric Approach to Sound Source Localization from Time-Delay Estimates. <i>IEEE/ACM Transactions on Audio Speech and Language Processing</i> , 2014, 22, 1082-1095.	5.8	64
32	RAVEL: an annotated corpus for training robots with audiovisual abilities. <i>Journal on Multimodal User Interfaces</i> , 2013, 7, 79-91.	2.9	27
33	Active-speaker detection and localization with microphones and cameras embedded into a robotic head. , 2013, , .		20