

Tatsuya Kawahara

List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/3681839/publications.pdf>

Version: 2024-02-01

234
papers

2,382
citations

471509

17
h-index

610901

24
g-index

237
all docs

237
docs citations

237
times ranked

963
citing authors

#	ARTICLE	IF	CITATIONS
1	ERICA: The ERATO Intelligent Conversational Android. , 2016, , .		91
2	Improved End-to-End Speech Emotion Recognition Using Self Attention Mechanism and Multitask Learning. , 0, , .		88
3	Statistical Speech Enhancement Based on Probabilistic Integration of Variational Autoencoder and Non-Negative Matrix Factorization. , 2018, , .		75
4	Flexible speech understanding based on combined key-phrase detection and verification. IEEE Transactions on Speech and Audio Processing, 1998, 6, 558-568.	1.5	56
5	A new phonetic tied-mixture model for efficient decoding. , 0, , .		48
6	Out-of-Domain Utterance Detection Using Classification Confidences of Multiple Topics. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 150-161.	3.2	37
7	Multilingual End-to-End Speech Translation. , 2019, , .		37
8	Language Model and Speaking Rate Adaptation for Spontaneous Presentation Speech Recognition. IEEE Transactions on Speech and Audio Processing, 2004, 12, 391-400.	1.5	36
9	Unsupervised Speech Enhancement Based on Multichannel NMF-Informed Beamforming for Noise-Robust Automatic Speech Recognition. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 960-971.	5.8	35
10	Fast Multichannel Nonnegative Matrix Factorization With Directivity-Aware Jointly-Diagonalizable Spatial Covariance Matrices for Blind Source Separation. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 2610-2625.	5.8	35
11	Prediction and Generation of Backchannel Form for Attentive Listening Systems. , 0, , .		35
12	Flexible mixed-initiative dialogue management using concept-level confidence measures of speech recognizer output. , 2000, , .		35
13	Talking with ERICA, an autonomous android. , 2016, , .		33
14	User Modeling in Spoken Dialogue Systems to Generate Flexible Guidance. User Modeling and User-Adapted Interaction, 2005, 15, 169-183.	3.8	32
15	Attentive listening system with backchanneling, response generation and flexible turn-taking. , 2017, , .		32
16	Acoustic-to-Word Attention-Based Model Complemented with Character-Level CTC-Based Model. , 2018, , .		31
17	Spoken Dialogue System for a Human-like Conversational Robot ERICA. Lecture Notes in Electrical Engineering, 2019, , 65-75.	0.4	31
18	A bootstrapping approach for developing language model of new spoken dialogue systems by selecting web texts. , 0, , .		29

#	ARTICLE	IF	CITATIONS
19	Bayesian Multichannel Speech Enhancement with a Deep Speech Prior. , 2018, , .		28
20	Semi-Supervised Multichannel Speech Enhancement With a Deep Speech Prior. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 2197-2212.	5.8	28
21	Gaussian mixture selection using context-independent HMM. , 0, , .		27
22	Bayesian Learning of a Language Model from Continuous Speech. IEICE Transactions on Information and Systems, 2012, E95-D, 614-625.	0.7	27
23	Leveraging Sequence-to-Sequence Speech Synthesis for Enhancing Acoustic-to-Word Speech Recognition. , 2018, , .		27
24	Multiparty Interaction Understanding Using Smart Multimodal Digital Signage. IEEE Transactions on Human-Machine Systems, 2014, 44, 625-637.	3.5	26
25	Computer Assisted Language Learning system based on dynamic question generation and error prediction for automatic speech recognition. Speech Communication, 2009, 51, 995-1005.	2.8	25
26	Practical use of English pronunciation system for Japanese students in the CALL classroom. , 0, , .		25
27	Robust Speech Recognition Based on Dereverberation Parameter Optimization Using Acoustic Model Likelihood. IEEE Transactions on Audio Speech and Language Processing, 2010, 18, 1708-1716.	3.2	24
28	Distilling the Knowledge of BERT for Sequence-to-Sequence ASR. , 0, , .		24
29	Automatic lecture transcription by exploiting presentation slide information for language model adaptation. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	23
30	Transfer Learning of Language-independent End-to-end ASR with Language Model Fusion. , 2019, , .		23
31	Japanese Dictation Toolkit. 1997 version.. Journal of the Acoustical Society of Japan (E), 1999, 20, 233-239.	0.1	23
32	End-to-End Speech Emotion Recognition Combined with Acoustic-to-Word ASR Model. , 0, , .		23
33	Prediction of Turn-taking Using Multitask Learning with Prediction of Backchannels and Fillers. , 0, , .		22
34	Speaker model selection based on the Bayesian information criterion applied to unsupervised speaker indexing. IEEE Transactions on Speech and Audio Processing, 2005, 13, 583-592.	1.5	21
35	Speech-Based Interactive Information Guidance System using Question-Answering Technique. , 2007, , .		21
36	Speech Enhancement Based on Bayesian Low-Rank and Sparse Decomposition of Multichannel Magnitude Spectrograms. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 215-230.	5.8	21

#	ARTICLE	IF	CITATIONS
37	Improving Transformer-Based Speech Recognition Systems with Compressed Structure and Speech Attributes Augmentation. , 0, , .		21
38	Smooth Turn-taking by a Robot Using an Online Continuous Model to Generate Turn-taking Cues. , 2019, , .		21
39	Conversational system for information navigation based on POMDP with user focus tracking. Computer Speech and Language, 2015, 34, 275-291.	4.3	20
40	Expressing reactive emotion based on multimodal emotion recognition for natural conversation in human-robot interaction. Advanced Robotics, 2019, 33, 1030-1041.	1.8	20
41	Automatic transcription of spontaneous lecture speech. , 0, , .		19
42	Bayes risk-based dialogue management for document retrieval system with speech interface. Speech Communication, 2010, 52, 61-71.	2.8	19
43	Independent Low-Rank Tensor Analysis for Audio Source Separation. , 2018, , .		19
44	Enhancing Monotonic Multihead Attention for Streaming ASR. , 0, , .		19
45	Cross-domain speech recognition using nonparallel corpora with cycle-consistent adversarial networks. , 2017, , .		18
46	Uyghur morpheme-based language models and ASR. , 2010, , .		17
47	Partial and synchronized captioning: A new tool to assist learners in developing second language listening skill. ReCALL, 2017, 29, 178-199.	5.2	17
48	Evaluation of Real-time Deep Learning Turn-taking Models for Multiple Dialogue Scenarios. , 2018, , .		17
49	Cross-Lingual Transfer Learning of Non-Native Acoustic Modeling for Pronunciation Error Detection and Diagnosis. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 391-401.	5.8	17
50	A Conversational Dialogue Manager for the Humanoid Robot ERICA. Lecture Notes in Electrical Engineering, 2019, , 119-131.	0.4	17
51	Joint Optimization of Denoising Autoencoder and DNN Acoustic Model Based on Multi-Target Learning for Noisy Speech Recognition. , 0, , .		17
52	An English pronunciation learning system for Japanese students based on diagnosis of critical pronunciation errors. ReCALL, 2004, 16, 173-188.	5.2	16
53	Multi-speaker Sequence-to-sequence Speech Synthesis for Data Augmentation in Acoustic-to-word Speech Recognition. , 2019, , .		16
54	Flexible guidance generation using user model in spoken dialogue systems. , 2003, , .		16

#	ARTICLE	IF	CITATIONS
55	Automatic Indexing of Lecture Presentations Using Unsupervised Learning of Presumed Discourse Markers. IEEE Transactions on Speech and Audio Processing, 2004, 12, 409-419.	1.5	15
56	A New ASR Evaluation Measure and Minimum Bayes-Risk Decoding for Open-domain Speech Understanding. , 0, , .		15
57	Statistical Transformation of Language and Pronunciation Models for Spontaneous Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2010, 18, 1539-1549.	3.2	15
58	Bayesian Multichannel Audio Source Separation Based on Integrated Source and Spatial Models. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 831-846.	5.8	15
59	Dialogue strategy to clarify user's queries for document retrieval system with speech interface. Speech Communication, 2006, 48, 1137-1150.	2.8	14
60	Multi-party human-robot interaction with distant-talking speech recognition. , 2012, , .		13
61	Emotion recognition by combining prosody and sentiment analysis for expressing reactive emotion by humanoid robot. , 2017, , .		13
62	Articulatory Modeling for Pronunciation Error Detection without Non-Native Training Data Based on DNN Transfer Learning. IEICE Transactions on Information and Systems, 2017, E100.D, 2174-2182.	0.7	13
63	Online Unsupervised Classification With Model Comparison in the Variational Bayes Framework for Voice Activity Detection. IEEE Journal on Selected Topics in Signal Processing, 2010, 4, 1071-1083.	10.8	12
64	Lexicon optimization based on discriminative learning for automatic speech recognition of agglutinative language. Speech Communication, 2014, 60, 78-87.	2.8	12
65	Named Entity Recognizer Trainable from Partially Annotated Data. Communications in Computer and Information Science, 2016, , 148-160.	0.5	12
66	Reverberant speech recognition combining deep neural networks and deep autoencoders augmented with a phone-class feature. Eurasip Journal on Advances in Signal Processing, 2015, 2015, .	1.7	11
67	Deep autoencoders augmented with phone-class feature for reverberant speech recognition. , 2015, , .		11
68	Exploiting automatic speech recognition errors to enhance partial and synchronized caption for facilitating second language listening. Computer Speech and Language, 2018, 49, 17-36.	4.3	11
69	Unsupervised Beamforming Based on Multichannel Nonnegative Matrix Factorization for Noisy Speech Recognition. , 2018, , .		11
70	Generalized Statistical Modeling of Pronunciation Variations using Variable-length Phone Context. , 0, , .		10
71	Language model transformation applied to lightly supervised training of acoustic model for congress meetings. , 2009, , .		10
72	Improved statistical models for SMT-based speaking style transformation. , 2010, , .		10

#	ARTICLE	IF	CITATIONS
73	ORTHROS: non-autoregressive end-to-end speech translation With dual-decoder. , 2021, , .		10
74	Latent Character Model for Engagement Recognition Based on Multimodal Behaviors. Lecture Notes in Electrical Engineering, 2019, , 119-130.	0.4	10
75	Topic-Independent Speaking-Style Transformation of Language Model for Spontaneous Speech Recognition. , 2007, , .		9
76	Construction of a Test Collection for Spoken Document Retrieval from Lecture Audio Data. Journal of Information Processing, 2009, 17, 82-94.	0.4	9
77	Optimizing spectral subtraction and wiener filtering for robust speech recognition in reverberant and noisy conditions. , 2010, , .		9
78	Generating a Variety of Backchannel Forms Based on Linguistic and Prosodic Features for Attentive Listening Agents. Transactions of the Japanese Society for Artificial Intelligence, 2016, 31, C-G31_1-10.	0.1	9
79	News Navigation System Based on Proactive Dialogue Strategy. , 2015, , 15-25.		9
80	Turn-Taking Prediction Based on Detection of Transition Relevance Place. , 0, , .		9
81	Multi-modal recording, analysis and indexing of poster sessions. , 0, , .		9
82	Automatic transcription system for meetings of the Japanese national congress. , 0, , .		9
83	Gaussian Mixture Optimization Based on Efficient Cross-Validation. IEEE Journal on Selected Topics in Signal Processing, 2010, 4, 540-547.	10.8	8
84	Morpheme concatenation approach in language modeling for large-vocabulary Uyghur speech recognition. , 2011, , .		8
85	Semi-Supervised Acoustic Model Training by Discriminative Data Selection From Multiple ASR Systemsâ€™ Hypotheses. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 1524-1534.	5.8	8
86	Effective articulatory modeling for pronunciation error detection of L2 learner without non-native training data. , 2017, , .		8
87	Semi-supervised ensemble DNN acoustic model training. , 2017, , .		8
88	Job Interviewer Android with Elaborate Follow-up Question Generation. , 2020, , .		8
89	Belief Network based Disambiguation of Object Reference in Spoken Dialogue System. Transactions of the Japanese Society for Artificial Intelligence, 2004, 19, 47-56.	0.1	8
90	Forward-Backward Attention Decoder. , 0, , .		8

#	ARTICLE	IF	CITATIONS
91	Efficient dialogue strategy to find users' intended items from information query results. , 2002, , .		8
92	Information Navigation System Based on POMDP that Tracks User Focus. , 2014, , .		8
93	Transcription and Distillation of Spontaneous Speech. , 2008, , 627-652.		7
94	GMM and HMM training by aggregated EM algorithm with increased ensemble sizes for robust parameter estimation. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	7
95	Modeling spoken decision support dialogue and optimization of its dialogue strategy. ACM Transactions on Speech and Language Processing, 2011, 7, 1-18.	0.9	7
96	Exploring deep neural networks and deep autoencoders in reverberant speech recognition. , 2014, , .		7
97	Engagement recognition by a latent character model based on multimodal listener behaviors in spoken dialogue. APSIPA Transactions on Signal and Information Processing, 2018, 7, .	3.3	7
98	Social Signal Detection in Spontaneous Dialogue Using Bidirectional LSTM-CTC. , 0, , .		7
99	Analysis of Effect and Timing of Fillers in Natural Turn-Taking. , 0, , .		7
100	Dialog navigator. , 2003, , .		7
101	Dependency structure analysis and sentence boundary detection in spontaneous Japanese. , 2004, , .		7
102	Analyzing temporal transition of real user's behaviors in a spoken dialogue system. , 0, , .		7
103	Optimization of dereverberation parameters based on likelihood of speech recognizer. , 0, , .		7
104	Estimation of interest and comprehension level of audience through multi-modal behaviors in poster conversations. , 0, , .		7
105	Generating Fillers Based on Dialog Act Pairs for Smooth Turn-Taking by Humanoid Robot. Lecture Notes in Electrical Engineering, 2019, , 91-101.	0.4	7
106	End-to-End Articulatory Attribute Modeling for Low-Resource Multilingual Speech Recognition. , 0, , .		7
107	Selective Multi-Task Learning For Speech Emotion Recognition Using Corpora Of Different Styles. , 2022, , .		7
108	Embedded Julius: Continuous Speech Recognition Software for Microprocessor. , 2006, , .		6

#	ARTICLE	IF	CITATIONS
109	Admissible stopping in viterbi beam search for unit selection in concatenative speech synthesis. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	6
110	New perspectives on spoken language understanding: Does machine need to fully understand speech?. , 2009, , .		6
111	A monotonic statistical machine translation approach to speaking style transformation. Computer Speech and Language, 2012, 26, 349-370.	4.3	6
112	Corpus and transcription system of Chinese Lecture Room. , 2014, , .		6
113	Annotation and analysis of listener's engagement based on multi-modal behaviors. , 2016, , .		6
114	Joint dialog act segmentation and recognition in human conversations using attention to dialog context. Computer Speech and Language, 2019, 57, 108-127.	4.3	6
115	Dialogue Speech Recognition by Combining Hierarchical Topic Classification and Language Model Switching. IEICE Transactions on Information and Systems, 2005, E88-D, 446-454.	0.7	6
116	Encoder Transfer for Attention-based Acoustic-to-word Speech Recognition. , 0, , .		6
117	Acoustic event detection for spotting "hot spots" in podcasts. , 0, , .		6
118	A WFST-based log-linear framework for speaking-style transformation. , 0, , .		6
119	Detection of hot spots in poster conversations based on reactive tokens of audience. , 0, , .		6
120	Smart Posterboard: Multi-modal Sensing and Analysis of Poster Conversations. , 2016, , 239-270.		6
121	Effective Prediction of Errors by Non-native Speakers Using Decision Tree for Speech Recognition-Based CALL System. IEICE Transactions on Information and Systems, 2009, E92-D, 2462-2468.	0.7	5
122	Hands-free human-robot communication robust to speaker's radial position. , 2013, , .		5
123	Unsupervised speaker adaptation of DNN-HMM by selecting similar speakers for lecture transcription. , 2014, , .		5
124	Synchrony in prosodic and linguistic features between backchannels and preceding utterances in attentive listening. , 2015, , .		5
125	Prediction of ice-breaking between participants using prosodic features in the first meeting dialogue. , 2016, , .		5
126	Pronunciation error detection using DNN articulatory model based on multi-lingual and multi-task learning. , 2016, , .		5

#	ARTICLE	IF	CITATIONS
127	Stem-Affix based Uyghur Morphological Analyzer. International Journal of Future Generation Communication and Networking, 2016, 9, 59-72.	0.7	5
128	Definition of Recipe Terms and Corpus Annotation for their Automatic Recognition. Journal of Natural Language Processing, 2015, 22, 107-131.	0.2	5
129	Generative Adversarial Training Data Adaptation for Very Low-Resource Automatic Speech Recognition. , 0, , .		5
130	Computationally-Efficient Overdetermined Blind Source Separation Based on Iterative Source Steering. IEEE Signal Processing Letters, 2022, 29, 927-931.	3.6	5
131	Introduction to the Special Issue on Spontaneous Speech Processing. IEEE Transactions on Speech and Audio Processing, 2004, 12, 349-350.	1.5	4
132	Effective error prediction using decision tree for ASR grammar network in call system. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	4
133	Using variational bayes free energy for unsupervised voice activity detection. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	4
134	Substring-based machine translation. Machine Translation, 2013, 27, 139-166.	1.3	4
135	Incorporating semantic information to selection of web texts for language model of spoken dialogue system. , 2013, , .		4
136	Optimized wavelet-domain filtering under noisy and reverberant conditions. APSIPA Transactions on Signal and Information Processing, 2015, 4, .	3.3	4
137	Assistive Typing Application for Older Adults Based on Input Stumble Detection. Journal of Information Processing, 2017, 25, 417-425.	0.4	4
138	Improving OOV Detection and Resolution with External Language Models in Acoustic-to-Word ASR. , 2018, , .		4
139	Improving CTC-based Acoustic Model with Very Deep Residual Time-delay Neural Networks. , 0, , .		4
140	Engagement Recognition in Spoken Dialogue via Neural Network by Aggregating Different Annotators' Models. , 0, , .		4
141	Dialogue Behavior Control Model for Expressing a Character of Humanoid Robots. , 2018, , .		4
142	Speech-based information retrieval system with clarification dialogue strategy. , 2005, , .		4
143	Language Model Adaptation Based on PLSA of Topics and Speakers for Automatic Transcription of Panel Discussions. IEICE Transactions on Information and Systems, 2005, E88-D, 439-445.	0.7	4
144	Predicting ASR errors by exploiting barge-in rate of individual users for spoken dialogue systems. , 0, , .		4

#	ARTICLE	IF	CITATIONS
145	Investigating Radical-Based End-to-End Speech Recognition Systems for Chinese Dialects and Japanese. , 0, , .		4
146	Khmer Speech Translation Corpus of the Extraordinary Chambers in the Courts of Cambodia (ECCC). , 2021, , .		4
147	HMM based on pair-wise Bayes classifiers. , 1992, , .		3
148	Automatic indexing of key sentences for lecture archives. , 0, , .		3
149	Smart posterboard: Multi-modal sensing and analysis of poster conversations. , 2013, , .		3
150	Speaker diarization based on audio-visual integration for smart posterboard. , 2014, , .		3
151	Language model adaptation for academic lectures using character recognition result of presentation slides. , 2015, , .		3
152	Multimodal interaction with the autonomous Android ERICA. , 2016, , .		3
153	Multi-lingual and multi-task DNN learning for articulatory error detection. , 2016, , .		3
154	Data selection from multiple ASR systems' hypotheses for unsupervised acoustic model training. , 2016, , .		3
155	Bayesian multichannel nonnegative matrix factorization for audio source separation and localization. , 2017, , .		3
156	Efficient Learning of Articulatory Models Based on Multi-Label Training and Label Correction for Pronunciation Learning. , 2018, , .		3
157	Multi-lingual Transformer Training for Khmer Automatic Speech Recognition. , 2019, , .		3
158	A Job Interview Dialogue System with Autonomous Android ERICA. Lecture Notes in Electrical Engineering, 2021, , 291-297.	0.4	3
159	Semi-supervised Multichannel Speech Separation Based on a Phone- and Speaker-Aware Deep Generative Model of Speech Spectrograms. , 2021, , .		3
160	Verification of Speech Recognition Results Incorporating In-domain Confidence and Discourse Coherence Measures. IEICE Transactions on Information and Systems, 2006, E89-D, 931-938.	0.7	3
161	Engagement Recognition from Listener's Behaviors in Spoken Dialogue Using a Latent Character Model. Transactions of the Japanese Society for Artificial Intelligence, 2018, 33, DSH-F_1-12.	0.1	3
162	ERICA and WikiTalk. , 2019, , .		3

#	ARTICLE	IF	CITATIONS
163	End-to-End Speech-to-Dialog-Act Recognition. , 0, , .		3
164	Autoregressive Moving Average Jointly-Diagonalizable Spatial Covariance Analysis for Joint Source Separation and Dereverberation. IEEE/ACM Transactions on Audio Speech and Language Processing, 2022, 30, 2368-2382.	5.8	3
165	Continuous speech recognition based on A* search with wordâ€pair constraint as heuristics. Systems and Computers in Japan, 1994, 25, 101-110.	0.2	2
166	Incorporating Dialogue Context and Topic Clustering in Out-of-Domain Detection. , 0, , .		2
167	Automatic Detection of Sentence and Clause Units using Local Syntactic Dependency. , 2007, , .		2
168	Real-Time Continuous Speech Recognition System on SH-4A Microprocessor. , 2007, , .		2
169	HMM training based on CV-EM and CV Gaussian mixture optimization. , 2007, , .		2
170	Voice Activity Detection Based on High Order Statistics and Online EM Algorithm. IEICE Transactions on Information and Systems, 2008, E91-D, 2854-2861.	0.7	2
171	Speech Activity Detection for Multi-Party Conversation Analyses Based on Likelihood Ratio Test on Spatial Magnitude. IEEE Transactions on Audio Speech and Language Processing, 2010, 18, 1354-1365.	3.2	2
172	Automatic transcription of parliamentary meetings and classroom lectures - A sustainable approach and real system evaluations -. , 2010, , .		2
173	Robust hands-free Automatic Speech Recognition for human-machine interaction. , 2010, , .		2
174	Discriminative approach to lexical entry selection for automatic speech recognition of agglutinative language. , 2012, , .		2
175	Semi-Blind speech enhancement based on recurrent neural network for source separation and dereverberation. , 2017, , .		2
176	Improving Very Deep Time-Delay Neural Network With Vertical-Attention For Effectively Training CTC-Based ASR Systems. , 2018, , .		2
177	Semi-autonomous avatar enabling unconstrained parallel conversations â€seamless hybrid of WOZ and autonomous dialogue systemsâ€. Advanced Robotics, 2021, 35, 657-663.	1.8	2
178	Multi-party Human-Machine Interaction Using a Smart Multimodal Digital Signage. Lecture Notes in Computer Science, 2013, , 408-415.	1.3	2
179	Automatic Transformation of Lecture Transcription into Document Style using Statistical Framework. Journal of Natural Language Processing, 2004, 11, 67-83.	0.2	2
180	A Model of Temporally Changing User Behaviors in a Deployed Spoken Dialogue System. Lecture Notes in Computer Science, 2009, , 409-414.	1.3	2

#	ARTICLE	IF	CITATIONS
181	Classroom note-taking system for hearing impaired students using automatic speech recognition adapted to lectures. , 0, , .		2
182	Evaluation Framework Design of Spoken Term Detection Study at the NTCIR-9 IR for Spoken Documents Task. Journal of Natural Language Processing, 2012, 19, 329-350.	0.2	2
183	Speaker diarization using eye-gaze information in multi-party conversations. , 0, , .		2
184	Interactional and pragmatics-related prosodic patterns in Mandarin dialog. , 0, , .		2
185	Combined Multi-Channel NMF-Based Robust Beamforming for Noisy Speech Recognition. , 0, , .		2
186	A Character Expression Model Affecting Spoken Dialogue Behaviors. Lecture Notes in Electrical Engineering, 2021, , 3-13.	0.4	2
187	Semi-Supervised Learning for Character Expression of Spoken Dialogue Systems. , 0, , .		2
188	CTC-Synchronous Training for Monotonic Attention Model. , 0, , .		2
189	Phone-Informed Refinement of Synthesized Mel Spectrogram for Data Augmentation in Speech Recognition. , 2022, , .		2
190	Speaker-independent consonant recognition by integrating discriminant analysis and hmm. Systems and Computers in Japan, 1991, 22, 84-95.	0.2	1
191	Spoken language processing for audio archives of lectures and panel discussions. , 0, , .		1
192	Unsupervised speaker indexing of discussions using anchor models. Systems and Computers in Japan, 2005, 36, 25-33.	0.2	1
193	Intelligent Transcription System Based on Spontaneous Speech Processing. , 2007, , .		1
194	Using online model comparison in the Variational Bayes framework for online unsupervised Voice Activity Detection. , 2010, , .		1
195	Probabilistic Concatenation Modeling for Corpus-Based Speech Synthesis. IEICE Transactions on Information and Systems, 2011, E94-D, 2006-2014.	0.7	1
196	Automatic classification of usability of ASR result for real-time captioning of lectures. , 2015, , .		1
197	Multi-modal sensing and analysis of poster conversations with smart posterboard. APSIPA Transactions on Signal and Information Processing, 2016, 5, .	3.3	1
198	Incremental training and constructing the very deep convolutional residual network acoustic models. , 2017, , .		1

#	ARTICLE	IF	CITATIONS
199	An End-to-End Approach to Joint Social Signal Detection and Automatic Speech Recognition. , 2018, , .		1
200	Voice Input Tutoring System for Older Adults using Input Stumble Detection. , 2018, , .		1
201	Audio-Visual Conversation Analysis by Smart Posterboard and Humanoid Robot. , 2018, , .		1
202	An Attentive Listening System for Autonomous Android ERICA: Comparative Evaluation with Human Attentive Listeners. Transactions of the Japanese Society for Artificial Intelligence, 2021, 36, H-L51_1-12.	0.1	1
203	An Interactive Framework for Document Retrieval and Presentation with Question-Answering Function in Restricted Domain. , 2007, , 126-134.		1
204	Recognition of Emotional States in Spoken Dialogue with a Robot. Lecture Notes in Computer Science, 2004, , 413-423.	1.3	1
205	Trigger-Based Language Model Adaptation for Automatic Transcription of Panel Discussions. IEICE Transactions on Information and Systems, 2006, E89-D, 1024-1031.	0.7	1
206	Bus Information System Based on User Models and Dynamic Generation of VoiceXML Scripts. Lecture Notes in Computer Science, 2007, , 46-60.	1.3	1
207	Recent Progress of Speech Recognition Technology. Journal of the Institute of Electrical Engineers of Japan, 2013, 133, 364-367.	0.0	1
208	Analysis and modeling of between-sentence pauses in news speech by Japanese newscasters. , 0, , .		1
209	Synthesizing waveform sequence-to-sequence to augment training data for sequence-to-sequence speech recognition. Acoustical Science and Technology, 2021, 42, 333-343.	0.5	1
210	Response Generation to Out-of-Database Questions for Example-Based Dialogue Systems. Lecture Notes in Electrical Engineering, 2021, , 303-317.	0.4	1
211	Individual identification by integrating facial image, walking image, and vocal features. Systems and Computers in Japan, 1997, 28, 77-84.	0.2	0
212	Embedded Julius on T-Engine Platform. , 2006, , .		0
213	Optimal learning of P-Layer additive F0 models with cross-validation. , 2009, , .		0
214	Japanese Painting Study Tool: A System for Creating Nihonga Portraits. , 2011, , .		0
215	Joint Phrase Alignment and Extraction for Statistical Machine Translation. Journal of Information Processing, 2012, 20, 512-523.	0.4	0
216	Admissible Stopping in Viterbi Beam Search for Unit Selection Speech Synthesis. IEICE Transactions on Information and Systems, 2013, E96.D, 1359-1367.	0.7	0

#	ARTICLE	IF	CITATIONS
217	Automatic Lecture Transcription Based on Discriminative Data Selection for Lightly Supervised Acoustic Model Training. IEICE Transactions on Information and Systems, 2015, E98.D, 1545-1552.	0.7	0
218	Confidence estimation for speech recognition systems using conditional random fields trained with partially annotated data. , 2016, , .		0
219	Automatic meeting transcription system for the Japanese parliament (diet). , 2017, , .		0
220	Engagement-Based Adaptive Behaviors for Laboratory Guide in Human-Robot Dialogue. Lecture Notes in Electrical Engineering, 2021, , 129-139.	0.4	0
221	Difference of acoustic modeling for read speech and dialogue speech.. Acoustical Science and Technology, 2001, 22, 373-374.	0.5	0
222	Formant structure estimation using vocal tract length normalization for CALL systems.. Acoustical Science and Technology, 2003, 24, 93-96.	0.5	0
223	Detection of quotations and inserted clauses and its application to dependency structure analysis in spontaneous Japanese. , 2006, , .		0
224	Spoken Dialogue System Based on Information Extraction from Web Text. Lecture Notes in Computer Science, 2010, , 196-197.	1.3	0
225	4. Captioning for Lectures Using Automatic Speech Recognition Technology. Kyokai Joho Imeji Zasshi/Journal of the Institute of Image Information and Television Engineers, 2012, 66, 641-644.	0.1	0
226	Language Modeling for Spoken Dialogue System Based on Sentence Filtering Using Predicate-Argument Structures. Transactions of the Japanese Society for Artificial Intelligence, 2014, 29, 53-59.	0.1	0
227	Comparison of discrete and continuous classifier-based HMM.. Journal of the Acoustical Society of Japan (E), 1992, 13, 361-367.	0.1	0
228	Managing Dialog and Joint Actions for Virtual Basketball Teammates. Lecture Notes in Computer Science, 2016, , 355-358.	1.3	0
229	Utterance Behavior of Users While Playing Basketball with a Virtual Teammate. , 2017, , .		0
230	Analysis of the Relationship Between Prosodic Features of Fillers and its Forms or Occurrence Positions. , 0, , .		0
231	Typing Tutor: Individualized Tutoring in Text Entry for Older Adults Based on Statistical Input Stumble Detection. Journal of Information Processing, 2018, 26, 362-374.	0.4	0
232	A Dialogue Behavior Control Model for Expressing a Characters of Humanoid Robots. Transactions of the Japanese Society for Artificial Intelligence, 2018, 33, C-137_1-9.	0.1	0
233	Autonomous Dialogue Technologies in Symbiotic Human-robot Interaction. , 2020, , .		0
234	Prediction of Shared Laughter for Human-Robot Dialogue. , 2020, , .		0