Lin-Shan Lee

List of Publications by Year
in descending order

151 papers	1,549 citations	567281 15 h-index	580821 25 g-index
152 all docs	152 docs citations	152 times ranked	534 citing authors

#	Article	IF	Citations
1	End-to-End Whispered Speech Recognition with Frequency-Weighted Approaches and Pseudo Whisper Pre-training. , 2021, , .		4
2	Defending Your Voice: Adversarial Attack on Voice Conversion. , 2021, , .		11
3	Fragmentvc: Any-To-Any Voice Conversion by End-To-End Extracting and Fusing Fine-Grained Voice Fragments with Attention. , 2021, , .		27
4	Towards Unsupervised Speech Recognition and Synthesis with Quantized Speech Representation Learning. , 2020, , .		18
5	Sequence-to-Sequence Automatic Speech Recognition with Word Embedding Regularization and Fused Decoding. , 2020, , .		7
6	Adversarial Training of End-to-end Speech Recognition Using a Criticizing Language Model. , 2019, , .		23
7	Unsupervised Discovery of Structured Acoustic Tokens With Applications to Spoken Term Detection. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 394-405.	5.8	5
8	Rhythm-Flexible Voice Conversion Without Parallel Data Using Cycle-GAN Over Phoneme Posteriorgram Sequences. , 2018, , .		13
9	Transcribing Lyrics from Commercial Song Audio: the First Step Towards Singing Content Processing. , 2018, , .		6
10	Scalable Sentiment for Sequence-to-Sequence Chatbot Response with Performance Analysis. , 2018, , .		13
11	Domain Independent Key Term Extraction from Spoken Content Based on Context and Term Location Information in the Utterances. , 2018, , .		1
12	Unsupervised Iterative Deep Learning of Speech Features and Acoustic Tokens with Applications to Spoken Term Detection. IEEE/ACM Transactions on Audio Speech and Language Processing, 2017, 25, 1914-1928.	5.8	7
13	Personalized acoustic modeling by weakly supervised multi-task deep learning using acoustic tokens discovered from unlabeled data. , 2017, , .		1
14	Abstractive headline generation for spoken content by attentive recurrent neural networks with ASR error modeling. , 2016, , .		4
15	Hierarchical attention model for improved machine comprehension of spoken content. , 2016, , .		8
16	Enhancing sparse voice annotation for semantic retrieval of personal photos by continuous space word representations. , 2015, , .		1
17	An iterative deep learning framework for unsupervised discovery of speech features and linguistic units with applications on spoken term detection. , 2015, , .		7
18	An Improved Framework for Recognizing Highly Imbalanced Bilingual Code-switched Lectures with Cross-language Acoustic Modeling and Frame-level Language Identification. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, , 1-1.	5.8	6

#	Article	IF	Citations
19	Supervised Detection and Unsupervised Discovery of Pronunciation Error Patterns for Computer-Assisted Language Learning. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, 23, 564-579.	5.8	24
20	Spoken Content Retrievalâ€"Beyond Cascading Speech Recognition with Text Retrieval. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, 23, 1389-1420.	5.8	73
21	Enhancing automatically discovered multi-level acoustic patterns considering context consistency with applications in spoken term detection. , 2015, , .		1
22	Unsupervised spoken term detection with spoken queries by multi-level acoustic patterns with varying model granularity. , 2014, , .		12
23	Towards personalized video summarization using synchronized comments and Probabilistic Latent Semantic Analysis. , 2014, , .		0
24	Improved Semantic Retrieval of Spoken Content by Document/Query Expansion with Random Walk Over Acoustic Similarity Graphs. IEEE/ACM Transactions on Audio Speech and Language Processing, 2014, 22, 80-94.	5.8	10
25	A Recursive Dialogue Game for Personalized Computer-Aided Pronunciation Training. IEEE/ACM Transactions on Audio Speech and Language Processing, 2014, , 1-1.	5.8	9
26	Spoken Knowledge Organization by Semantic Structuring and a Prototype Course Lecture System for Personalized Learning. IEEE/ACM Transactions on Audio Speech and Language Processing, 2014, 22, 883-898.	5.8	14
27	Transcribing code-switched bilingual lectures using deep neural networks with unit merging in acoustic modeling. , 2014, , .		4
28	Improved open-vocabulary spoken content retrieval with word and subword lattices using acoustic feature similarity. Computer Speech and Language, 2014, 28, 1045-1065.	4.3	7
29	Variable Selection Linear Regression for Robust Speech Recognition. IEICE Transactions on Information and Systems, 2014, E97.D, 1477-1487.	0.7	1
30	Personalized video summarization based on Multi-Layered Probabilistic Latent Semantic Analysis with shared topics. , 2014, , .		2
31	An Experimental Analysis on Integrating Multi-Stream Spectro-Temporal, Cepstral and Pitch Information for Mandarin Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2013, 21, 2006-2014.	3.2	9
32	Enhanced Spoken Term Detection Using Support Vector Machines and Weighted Pseudo Examples. IEEE Transactions on Audio Speech and Language Processing, 2013, 21, 1272-1284.	3.2	17
33	Model-Based Unsupervised Spoken Term Detection with Spoken Queries. IEEE Transactions on Audio Speech and Language Processing, 2013, 21, 1330-1342.	3.2	22
34	Towards unsupervised semantic retrieval of spoken content with query expansion based on automatically discovered acoustic patterns. , 2013, , .		6
35	Toward unsupervised discovery of pronunciation error patterns using universal phoneme posteriorgram for computer-assisted language learning. , 2013, , .		16
36	Unsupervised domain adaptation for spoken document summarization with structured support vector machine. , 2013, , .		5

#	Article	IF	Citations
37	A dialogue game framework with personalized training using reinforcement learning for computer-assisted language learning. , 2013, , .		13
38	Unsupervised discovery of linguistic structure including two-level acoustic patterns using three cascaded stages of iterative optimization. , 2013, , .		17
39	Enhancing query expansion for semantic retrieval of spoken content with automatically discovered acoustic patterns. , 2013, , .		10
40	Utterance-level latent topic transition modeling for spoken documents and its application in automatic summarization. , 2012, , .		3
41	Two-dimensional frame-and-feature weighted Viterbi decoding for robust speech recognition. , 2012, , .		0
42	Recognition of highly imbalanced code-mixed bilingual speech with frame-level language detection based on blurred posteriorgram. , 2012, , .		7
43	Minimum Phone Error model training on merged acoustic units for transcribing bilingual code-switched speech. , 2012, , .		2
44	Semantic query expansion and context-based discriminative term modeling for spoken document retrieval. , 2012, , .		4
45	Improved approaches of modeling and detecting Error Patterns with empirical analysis for Computer-Aided Pronunciation Training. , 2012, , .		34
46	Improved semantic retrieval of spoken content by language models enhanced with acoustic similarity graph. , 2012, , .		8
47	Personalized language modeling by crowd sourcing with social network data for voice access of cloud applications. , 2012, , .		9
48	Interactive Spoken Document Retrieval With Suggested Key Terms Ranked by a Markov Decision Process. IEEE Transactions on Audio Speech and Language Processing, 2012, 20, 632-645.	3.2	21
49	Modulation Spectrum Equalization for Improved Robust Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2012, 20, 828-843.	3.2	19
50	Integrating Recognition and Retrieval With Relevance Feedback for Spoken Term Detection. IEEE Transactions on Audio Speech and Language Processing, 2012, 20, 2095-2110.	3.2	9
51	Semantic Analysis and Organization of Spoken Documents Based on Parameters Derived From Latent Topics. IEEE Transactions on Audio Speech and Language Processing, 2011, 19, 1875-1889.	3.2	15
52	Integrating frame-based and segment-based dynamic time warping for unsupervised spoken term detection with spoken queries. , 2011, , .		18
53	Improved spoken term detection with graph-based re-ranking in feature space. , 2011, , .		17
54	Multi-stream spectro-temporal and cepstral features based on data-driven hierarchical phoneme clusters. , 2011, , .		1

#	Article	IF	Citations
55	Improved spoken term detection using support vector machines based on lattice context consistency. , 2011, , .		9
56	Performance Analysis for Lattice-Based Speech Indexing Approaches Using Words and Subword Units. IEEE Transactions on Audio Speech and Language Processing, 2010, 18, 1562-1574.	3.2	17
57	An initial attempt to improve spoken term detection by learning optimal weights for different indexing features. , 2010, , .		2
58	Integrating recognition and retrieval with user feedback: A new framework for spoken term detection. , 2010, , .		6
59	A framework integrating different relevance feedback scenarios and approaches for spoken term detection. , 2010, , .		3
60	Automatic key term extraction from spoken course lectures using branching entropy and prosodic/semantic features. , 2010, , .		22
61	Latent semantic retrieval of personal photos with sparse user annotation by fused image/speech/text features. , 2009, , .		5
62	Improved lattice-based spoken document retrieval by directly learning from the evaluation measures. , 2009, , .		10
63	Improved clustered hierarchical tandem system with bottom-up processing. , 2009, , .		2
64	Higher Order Cepstral Moment Normalization for Improved Robust Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 205-220.	3.2	24
65	Improved Features and Models for Detecting Edit Disfluencies in Transcribing Spontaneous Mandarin Speech. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 1263-1278.	3.2	14
66	Voice-based information retrieval &#x2014; how far are we from the text-based information retrieval ?. , 2009, , .		8
67	Histogram-Based Quantization for Robust and/or Distributed Speech Recognition. IEEE Transactions on Audio Speech and Language Processing, 2008, 16, 859-873.	3.2	12
68	Latent semantic retrieval of spoken documents over position specific posterior lattices. , 2008, , .		9
69	Evaluation and Analysis of Minimum Phone Error Training and its Modified Versions for Large Vocabulary Mandarin Speech Recognition. , 2008, , .		0
70	Context dependent quantization for distributed and/or robust speech recognition. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	0
71	Type-II dialogue systems for information access from unstructured knowledge sources. , 2007, , .		3
72	Analytical comparison between position specific posterior lattices and confusion networks based on words and subword units for spoken document indexing. , 2007, , .		12

#	Article	IF	Citations
73	Robust topic inference for latent semantic language model adaptation. , 2007, , .		10
74	Robust speech recognition by properly utilizing reliable frames and segments in corrupted signals. , 2007, , .		1
75	Modulation spectrum equalization for robust speech recognition. , 2007, , .		4
76	Three-Stage Error Concealment for Distributed Speech Recognition (DSR) with Histogram-Based Quantization (HQ) Under Noisy Environment. , 2007, , .		2
77	Virtual Conduction System with Multi-Resolution Wall Display. , 2007, , .		0
78	Pronunciation Modeling With Reduced Confusion for Mandarin Chinese Using a Three-Stage Framework. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 661-675.	3.2	9
79	A Perceptually Constrained GSVD-Based Approach for Enhancing Speech Corrupted by Colored Noise. IEEE Transactions on Audio Speech and Language Processing, 2007, 15, 119-134.	3.2	12
80	A Multi-layered Summarization System for Multi-media Archives by Understanding and Structuring of Chinese Spoken Documents. Lecture Notes in Computer Science, 2006, , 683-692.	1.3	1
81	Named entity recognition from spoken documents using global evidences and external knowledge sources with applications on Mandarin Chinese. , 2005, , .		6
82	A multi-modal dialogue system for information navigation and retrieval across spoken document archives with topic hierarchies. , 2005, , .		5
83	Segmental eigenvoice with delicate eigenspace for improved speaker adaptation. IEEE Transactions on Speech and Audio Processing, 2005, 13, 399-411.	1.5	14
84	Spoken document understanding and organization. IEEE Signal Processing Magazine, 2005, 22, 42-60.	5.6	111
85	A discriminative HMM/N-gram-based retrieval approach for mandarin spoken documents. ACM Transactions on Asian Language Information Processing, 2004, 3, 128-145.	0.8	26
86	Data-driven temporal filters for robust features in speech recognition obtained via Minimum Classification Error (MCE). , 2002, , .		9
87	Discriminating capabilities of syllable-based features and approaches of utilizing them for voice retrieval of speech information in Mandarin Chinese. IEEE Transactions on Speech and Audio Processing, 2002, 10, 303-314.	1.5	45
88	A set of corpus-based text-to-speech synthesis technologies for Mandarin Chinese. IEEE Transactions on Speech and Audio Processing, 2002, 10, 481-494.	1.5	27
89	A hierarchical tag-graph search scheme with layered grammar rules for spontaneous speech understanding. Pattern Recognition Letters, 2002, 23, 819-831.	4.2	6
90	Voice access of global information for broad-band wireless: technologies of today and challenges of tomorrow. Proceedings of the IEEE, 2001, 89, 41-57.	21.3	8

#	Article	IF	Citations
91	Improved parallel model combination techniques with split Gaussian mixtures for speech recognition under noisy conditions. , 1999, , .		1
92	A framework of performance evaluation and error analysis methodology for speech understanding systems. , 1999, , .		0
93	Intelligent retrieval of dynamic networked information from mobile terminals using spoken natural language queries. IEEE Transactions on Consumer Electronics, 1998, 44, 62-72.	3.6	3
94	Isolated Mandarin base-syllable recognition based upon the segmental probability model. IEEE Transactions on Speech and Audio Processing, 1998, 6, 293-299.	1.5	4
95	Complete recognition of continuous Mandarin speech for Chinese language with very large vocabulary using limited training data. IEEE Transactions on Speech and Audio Processing, 1997, 5, 195-200.	1.5	38
96	Isolated Mandarin syllable recognition with limited training data specially considering the effect of tones. IEEE Transactions on Speech and Audio Processing, 1997, 5, 75-80.	1.5	5
97	Voice dictation of Mandarin Chinese. IEEE Signal Processing Magazine, 1997, 14, 63-101.	5.6	54
98	Continuous hidden Markov models integrating transitional and instantaneous features for Mandarin syllable recognition. Computer Speech and Language, 1993, 7, 247-263.	4.3	14
99	Markov modeling of Mandarin Chinese for decoding the phonetic sequence into Chinese characters. Computer Speech and Language, 1991, 5, 363-377.	4.3	21
100	Special speech recognition approaches for the highly confusing Mandarin syllables based on hidden Markov models. Computer Speech and Language, 1991, 5, 181-201.	4.3	10
101	The synthesis rules in a Chinese text-to-speech system. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989, 37, 1309-1320.	2.0	61
102	The Minimum Likelihood--A New Concept for Bit Synchronization. IEEE Transactions on Communications, 1987, 35, 545-549.	7.8	12
103	A General Theory for Asynchronous Speech Encryption Techniques. IEEE Journal on Selected Areas in Communications, 1986, 4, 280-287.	14.0	6
104	A Differential Technique for Detections of Circularly Polarized Satellite Signal Parameters. IEEE Transactions on Communications, 1984, 32, 1135-1139.	7.8	1
105	A New Formulation of Spectrum-Orbit Utilization Efficiency for Satellite Communications in Interference-Limited Situations. IRE Transactions on Communications Systems, 1984, 32, 212-214.	0.6	0
106	A New Frequency Domain Speech Scrambling System Which Does Not Require Frame Synchronization. IEEE Transactions on Communications, 1984, 32, 444-456.	7.8	28
107	An A, B-Polar Approach for Multimode Polarization Analysis in Satellite Communications. IRE Transactions on Communications Systems, 1984, 32, 1045-1049.	0.6	0
108	A New Time Domain Speech Scrambling System Which Does Not Require Frame Synchronization. IEEE Journal on Selected Areas in Communications, 1984, 2, 443-455.	14.0	7

#	Article	IF	Citations
109	Asynchronous speech encryption - formulation and simulation. , 1983, , .		0
110	Accurate Detection of Satellite Beacon Polarization States Using Cascaded Heterodyne Phase-Locked Loops. IEEE Transactions on Communications, 1982, 30, 283-293.	7.8	1
111	A modified blind equalization technique based on a constant modulus algorithm. , 0, , .		5
112	A Chinese text-to-speech system based upon a syllable concatenation model. , 0, , .		3
113	An Exact Performance Analysis of the Linear Diversity Combining Receiver for Fh/mfsk Systems Against Band Multitone Jammer. , 0, , .		0
114	An improved pragmatic approach to coded modulation design with multiple carrier frequencies. , 0, , .		0
115	Panel 1: Are We Providing The Right Education For Computer Science/engineering Students?. , 0, , .		0
116	Golden Mandarin (III)-a user-adaptive prosodic-segment-based Mandarin dictation machine for Chinese language with very large vocabulary. , 0, , .		11
117	Complete recognition of continuous Mandarin speech for Chinese language with very large vocabulary but limited training data. , 0, , .		18
118	Speaker intention modeling for large vocabulary Mandarin spoken dialogues. , 0, , .		4
119	Robust speech recognition features based on temporal trajectory filtering of frequency band spectrum. , 0, , .		11
120	Very-large-vocabulary Mandarin voice message file retrieval using speech queries. , 0, , .		4
121	An efficient voice retrieval system for very-large-vocabulary Chinese textual databases with a clustered language model. , 0, , .		5
122	Automatic generation of prosodic structure for high quality Mandarin speech synthesis. , 0, , .		15
123	Use of prosodic information to integrate acoustic and linguistic knowledge in continuous Mandarin speech recognition with very large vocabulary. , 0, , .		6
124	Fast and accurate recognition of very-large-vocabulary continuous Mandarin speech for Chinese language with improved segmental probability modeling. , 0, , .		3
125	A modified code tracking loop for direct-sequence spread-spectrum systems on frequency-selective fading channels. , 0, , .		1
126	Internet Chinese information retrieval using unconstrained Mandarin speech queries based on a client-server architecture and a PAT-tree-based language model. , 0, , .		2

#	ARTICLE	IF	CITATIONS
127	Syllable-based relevance feedback techniques for Mandarin voice record retrieval using speech queries. , 0, , .		1
128	A key-phrase understanding framework integrating real world knowledge with speech recognition with initial application in voice memo systems for Mandarin Chinese. , 0, , .		2
129	A Chinese text-to-speech system based on part-of-speech analysis, prosodic modeling and non-uniform units. , 0, , .		16
130	A multi-phase approach for fast spotting of large vocabulary Chinese keywords from Mandarin speech using prosodic information. , 0, , .		6
131	A word-length-dependent confidence measure for large vocabulary Chinese keyword spotting. , 0, , .		0
132	Improved robustness for speech recognition under noisy conditions using correlated parallel model combination. , 0, , .		4
133	Statistics-based segment pattern lexicon-a new direction for Chinese language modeling. , 0, , .		9
134	A noncoherent sequential PN code acquisition scheme in DS/SS receivers. , 0, , .		1
135	Fundamental performance analysis for spoken dialogue systems based on a quantitative simulation approach. , 0, , .		1
136	Retrieval of broadcast news speech in Mandarin Chinese collected in Taiwan using syllable-level statistical characteristics. , 0, , .		23
137	Improved pronunciation modelling by inverse word frequency and pronunciation entropy. , 0, , .		1
138	Improved MFCC feature extraction by PCA-optimized filter-bank for speech recognition. , 0, , .		13
139	Rapid speaker adaptation using a priori knowledge by eigenspace analysis of MLLR parameters. , 0, , .		14
140	Pronunciation variation analysis based on acoustic and phonemic distance measures with application examples on Mandarin Chinese. , 0, , .		5
141	Improved robust features for speech recognition by integrating time-frequency principal components (TFPC) and histogram equalization (HEQ). , 0, , .		3
142	Improved tone recognition for fluent Mandarin speech based on new inter-syllabic features and robust pitch extraction. , 0, , .		4
143	Improved language model adaptation using existing and derived external resources. , 0, , .		2
144	Data-driven temporal filters based on multi-eigenvectors for robust features in speech recognition. , 0, , .		5

#	ARTICLE	IF	CITATIONS
145	Large vocabulary continuous Mandarin speech recognition using finite state machine. , 0, , .		0
146	Efficient and robust distributed speech recognition (DSR) over wireless fading channels: 2D-DCT compression, iterative bit allocation, short BCH code and interleaving. , 0, , .		10
147	Robust features for speech recognition using minimum variance distortionless response (MVDR) spectrum estimation and feature normalization techniques. , 0, , .		3
148	Joint Uncertainty Decoding (JUD) with Histogram-Based Quantization (HQ) for Robust and/or Distributed Speech Recognition. , 0, , .		3
149	Entropy-Based Feature Parameter Weighting for Robust Speech Recognition. , 0, , .		4
150	Improved Spoken Document Retrieval With Dynamic Key Term Lexicon and Probabilistic Latent Semantic Analysis (PLSA). , 0, , .		7
151	Improved search strategy for large vocabulary continuous Mandarin speech recognition. , 0, , .		2