Engin Erzin

List of Publications by Year
in descending order

99 papers	1,297 citations	567281 15 h-index	454955 30 g-index
99 all docs	99 docs citations	99 times ranked	942 citing authors

#	ARTICLE	IF	CITATIONS
1	Audiovisual Synchronization and Fusion Using Canonical Correlation Analysis. IEEE Transactions on Multimedia, 2007, 9, 1396-1403.	7.2	144
2	Discriminative Analysis of Lip Motion Features for Speaker Identification and Speech-Reading. IEEE Transactions on Image Processing, 2006, 15, 2879-2891.	9.8	100
3	Teager energy based feature parameters for speech recognition in car noise. IEEE Signal Processing Letters, 1999, 6, 259-261.	3.6	96
4	Adaptive filtering for non-Gaussian stable processes. IEEE Signal Processing Letters, 1994, 1, 163-165.	3.6	93
5	Multimodal speaker identification using an adaptive classifier cascade based on modality reliability. IEEE Transactions on Multimedia, 2005, 7, 840-852.	7.2	60
6	Analysis of Head Gesture and Prosody Patterns for Prosody-Driven Head-Gesture Animation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30, 1330-1345.	13.9	59
7	Multimodal speaker/speech recognition using lip motion, lip texture and audio. Signal Processing, 2006, 86, 3549-3558.	3.7	50
8	Formant position based weighted spectral features for emotion recognition. Speech Communication, 2011, 53, 1186-1197.	2.8	49
9	Multimodal Person Recognition for Human-Vehicle Interaction. IEEE MultiMedia, 2006, 13, 18-31.	1.7	47
10	Learn2Dance: Learning Statistical Music-to-Dance Mappings for Choreography Synthesis. IEEE Transactions on Multimedia, 2012, 14, 747-759.	7.2	42
11	Improving Throat Microphone Speech Recognition by Joint Analysis of Throat and Acoustic Microphone Recordings. IEEE Transactions on Audio Speech and Language Processing, 2009, 17, 1316-1324.	3.2	35
12	Combined Gesture-Speech Analysis and Speech Driven Gesture Synthesis. , 2006, , .		30
13	Source and Filter Estimation for Throat-Microphone Speech Enhancement. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 265-275.	5.8	30
14	Artificial bandwidth extension of spectral envelope along a Viterbi path. Speech Communication, 2013, 55, 111-118.	2.8	27
15	Multimodal analysis of speech and arm motion for prosody-driven synthesis of beat gestures. Speech Communication, 2016, 85, 29-42.	2.8	23
16	Prosody-Driven Head-Gesture Animation. , 2007, , .		22
17	Automatic Classification of Musical Genres Using Inter-Genre Similarity. IEEE Signal Processing Letters, 2007, 14, 521-524.	3.6	21
18	Analysis of Engagement and User Experience with a Laughter Responsive Social Robot. , 0, , .		19

#	Article	IF	Citations
19	Use of Line Spectral Frequencies for Emotion Recognition from Speech. , 2010, , .		15
20	Analysis of interaction attitudes using data-driven hand gesture phrases. , 2014, , .		15
21	Robust Lip-Motion Features For Speaker Identification. , 0, , .		12
22	Comparison of Phoneme and Viseme Based Acoustic Units for Speech Driven Realistic Lip Animation. , 2007, , .		12
23	An audio-driven dancing avatar. Journal on Multimodal User Interfaces, 2008, 2, 93-103.	2.9	12
24	Improving phoneme recognition of throat microphone speech recordings using transfer learning. Speech Communication, 2021, 129, 25-32.	2.8	12
25	Driver Status Identification from Driving Behavior Signals. , 2012, , 31-55.		11
26	Monitoring Infant's Emotional Cry in Domestic Environments Using the Capsule Network Architecture. , 0, , .		11
27	Discriminative lip-motion features for biometric speaker identification. , 0, , .		10
28	A subjective listening test of six different artificial bandwidth extension approaches in English, Chinese, German, and Korean. , 2016, , .		10
29	Multifaceted Engagement in Social Interaction with a Machine: The JOKER Project. , 2018, , .		10
30	Vocal Tract Airway Tissue Boundary Tracking for rtMRI Using Shape and Appearance Priors. , 0, , .		10
31	Comparison of Phoneme and Viseme Based Acoustic Units for Speech Driven Realistic lip Animation. , 2007, , .		9
32	Audio-Facial Laughter Detection in Naturalistic Dyadic Conversations. IEEE Transactions on Affective Computing, 2017, 8, 534-545.	8.3	9
33	Multimodal analysis of speech prosody and upper body gestures using hidden semi-Markov models. , 2013, , .		8
34	Affect-expressive hand gestures synthesis and animation. , 2015, , .		8
35	Affective synthesis and animation of arm gestures from speech prosody. Speech Communication, 2020, 119, 1-11.	2.8	8
36	Estimation and Analysis of Facial Animation Parameter Patterns. , 2007, , .		7

#	Article	IF	Citations
37	Empirical Mode Decomposition of Throat Microphone Recordings for Intake Classification. , 2017, , .		7
38	Batch Recurrent Q-Learning for Backchannel Generation Towards Engaging Agents. , 2019, , .		7
39	Continuous emotion tracking using total variability space. , 0, , .		7
40	Interframe differential coding of line spectrum frequencies. IEEE Transactions on Speech and Audio Processing, 1994, 2, 350-352.	1.5	6
41	Joint Audio-Video Processing for Robust Biometric Speaker Identification in Car. , 2005, , 237-256.		6
42	Multicamera Audio-Visual Analysis of Dance Figures. , 2007, , .		6
43	Audio-driven human body motion analysis and synthesis. Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2008, , .	1.8	6
44	INTERSPEECH 2009 Emotion Recognition Challenge evaluation. , 2010, , .		6
45	The JESTKOD database: an affective multimodal database of dyadic interactions. Language Resources and Evaluation, 2017, 51, 857-872.	2.7	6
46	Audio-Visual Prediction of Head-Nod and Turn-Taking Events in Dyadic Interactions. , 0, , .		6
47	Vocal Tract Contour Tracking in rtMRI Using Deep Temporal Regression Network. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 3053-3064.	5.8	6
48	Training Socially Engaging Robots: Modeling Backchannel Behaviors with Batch Reinforcement Learning. IEEE Transactions on Affective Computing, 2022, 13, 1840-1853.	8.3	6
49	Speech-Driven Automatic Facial Expression Synthesis. , 2008, , .		5
50	Use of Agreement/Disagreement Classification in Dyadic Interactions for Continuous Emotion Recognition. , 0, , .		5
51	Multi-modal analysis of dance performances for music-driven choreography synthesis. , 2010, , .		4
52	Detection of Food Intake Events From Throat Microphone Recordings Using Convolutional Neural Networks. , 2018, , .		4
53	Shaped fixed codebook search for CELP coding at low bit rates. , 0, , .		3
54	Unsupervised dance figure analysis from video for dancing Avatar animation. , 2008, , .		3

#	Article	IF	Citations
55	Artificial bandwidth extension of spectral envelope with temporal clustering. , 2011, , .		3
56	Evaluation of emotion recognition from speech. , 2012, , .		3
57	Enhancement of throat microphone recordings by learning phone-dependent mappings of speech spectra. , 2013, , .		3
58	JESTKOD database: Dyadic interaction analysis. , 2015, , .		3
59	Affect burst detection using multi-modal cues. , 2015, , .		3
60	Food intake classification using throat microphone. , 2016, , .		3
61	Agreement and disagreement classification of dyadic interactions using vocal and gestural cues. , 2016, , .		3
62	Speech features for telemonitoring of Parkinson's disease symptoms. , 2017, 2017, 3801-3805.		3
63	A Deep Learning Approach for Data Driven Vocal Tract Area Function Estimation. , 2018, , .		3
64	Emotion Dependent Domain Adaptation for Speech Driven Affective Facial Feature Synthesis. IEEE Transactions on Affective Computing, 2022, 13, 1501-1513.	8.3	3
65	RANSAC-Based Training Data Selection on Spectral Features for Emotion Recognition from Spontaneous Speech. Lecture Notes in Computer Science, 2011, , 36-47.	1.3	3
66	Cross-Subject Continuous Emotion Recognition Using Speech and Body Motion in Dyadic Interactions. , 0, , .		3
67	Estimation of Acoustic Microphone Vocal Tract Parameters from Throat Microphone Recordings. , 2009, , 161-169.		2
68	Automatic emotion recognition for facial expression animation from speech. , 2009, , .		2
69	Affect burst recognition using multi-modal cues. , 2014, , .		2
70	On the importance of hidden bias and hidden entropy in representational efficiency of the Gaussian-Bipolar Restricted Boltzmann Machines. Neural Networks, 2018, 105, 405-418.	5.9	2
71	Food intake detection using autoencoder-based deep neural networks. , 2018, , .		2
72	Head Nod Detection in Dyadic Conversations. , 2019, , .		2

#	Article	IF	Citations
73	Extracting Gene Regulation Information from Microarray Time-Series Data Using Hidden Markov Models. Lecture Notes in Computer Science, 2006, , 144-153.	1.3	2
74	A new statistical excitation mapping for enhancement of throat microphone recordings. , 0, , .		2
75	Inter Genre Similarity Modeling For Automatic Music Genre Classification. , 0, , .		1
76	Boosting Classifiers for Music Genre Classification. , 0, , .		1
77	Analysis of JestKOD database using affective state annotations. , 2016, , .		1
78	Real-time speech driven gesture animation. , 2016, , .		1
79	Real-time audiovisual laughter detection. , 2017, , .		1
80	A New Interface for Affective State Estimation and Annotation from Speech. , 2019, , .		1
81	A Diversity Combination Model Incorporating an Inward Bias for Interaural Time-Level Difference Cue Integration in Sound Lateralization. Applied Sciences (Switzerland), 2020, 10, 6356.	2.5	1
82	Use of affect context in dyadic interactions for continuous emotion recognition. Speech Communication, 2021, 132, 70-82.	2.8	1
83	Engagement Rewarded Actor-Critic with Conservative Q-Learning for Speech-Driven Laughter Backchannel Generation. , 2021, , .		1
84	AffectON: Incorporating Affect Into Dialog Generation. IEEE Transactions on Affective Computing, 2023, 14, 823-835.	8.3	1
85	Line spectral frequency representation of subbands for speech recognition. Signal Processing, 1995, 44, 117-119.	3.7	0
86	Speech Driven 3D Head Gesture Synthesis. , 0, , .		0
87	Learning Gene Regulation from Microarray Data via Hidden Markov Models. , 2007, , .		0
88	Joint Correlation Analysis of Audio-Visual Dance Figures. , 2007, , .		0
89	Estimation of Personalized Facial Gesture Patterns. , 2007, , .		0
90	Analysis and synthesis of multiview audio-visual dance figures. , 2008, , .		0

#	Article	IF	Citations
91	Correction to "Learn2Dance: Learning Statistical Music-to-Dance Mappings for Choreography Synthesis" [Jun 12 747-759]. IEEE Transactions on Multimedia, 2012, 14, 1376-1376.	7.2	0
92	Speech rhythm-driven gesture animation. , 2013, , .		0
93	A phonetic classification for throat microphone enhancement. , 2014, , .		0
94	Artificial bandwidth extension of speech excitation. , 2015, , .		0
95	Multimodal prediction of head nods in dyadic conversations. , 2018, , .		0
96	Use of non-verbal vocalizations for continuous emotion recognition from speech and head motion. , 2019, , .		0
97	Automatic Vocal Tractlandmark Tracking in Rtmri Using Fully Convolutional Networks and Kalman Filter. , 2020, , .		0
98	Emotion Dependent Facial Animation from Affective Speech. , 2020, , .		0
99	Domain Adaptation for Food Intake Classification with Teacher/Student Learning. IEEE Transactions on Multimedia, 2020, , 1-1.	7.2	0