

IEEE TRANSACTIONS ON *MULTIMEDIA*

A PUBLICATION OF
THE IEEE CIRCUITS AND SYSTEMS SOCIETY
THE IEEE SIGNAL PROCESSING SOCIETY
THE IEEE COMMUNICATIONS SOCIETY
THE IEEE COMPUTER SOCIETY



<http://www.signalprocessingsociety.org/tmm/>

JUNE 2020

VOLUME 22

NUMBER 6

ITMUF8

(ISSN 1520-9210)

REGULAR PAPERS

Audio/Speech/Language Analysis and Synthesis

Acoustic Scene Clustering Using Joint Optimization of Deep Embedding Learning and Clustering Iteration *Y. Li, M. Liu, W. Wang, Y. Zhang, and Q. He* 1385

Compression and Coding

Rate Constrained Multiple-QP Optimization for HEVC *M. Wang, J. Xiong, L. Xu, W. Xie, K. N. Ngan, and J. Qin* 1395

Image/Video/Graphics Analysis and Synthesis

A Single-Image Super-Resolution Method Based on Progressive-Iterative Approximation *Y. Zhang, P. Wang, F. Bao, X. Yao, C. Zhang, and H. Lin* 1407

Interact as You Intend: Intention-Driven Human-Object Interaction Detection *B. Xu, J. Li, Y. Wong, Q. Zhao, and M. S. Kankanhalli* 1423

2D Pose-Based Real-Time Human Action Recognition With Occlusion-Handling *F. Angelini, Z. Fu, Y. Long, L. Shao, and S. M. Naqvi* 1433

Light Field Super-Resolution Using Edge-Preserved Graph-Based Regularization *V. K. Ghassab and N. Bouguila* 1447

WeGAN: Deep Image Hashing With Weighted Generative Adversarial Networks *Y. Wang, L. Zhang, F. Nie, X. Li, Z. Chen, and F. Wang* 1458

Sparse Multimedia Signal Processing

Multi-Direction Dictionary Learning Based Depth Map Super-Resolution With Autoregressive Modeling *J. Wang, W. Xu, J.-F. Cai, Q. Zhu, Y. Shi, and B. Yin* 1470

Display Technology for Multimedia

Adaptive Single Image Dehazing Using Joint Local-Global Illumination Adjustment *H.-M. Hu, H. Zhang, Z. Zhao, B. Li, and J. Zheng* 1485

(Contents Continued on Back Cover)



<i>Multimedia Search and Retrieval</i>	
Multi-View Saliency Guided Deep Neural Network for 3-D Object Retrieval and Classification	1496
. <i>H.-Y. Zhou, A.-A. Liu, W.-Z. Nie, and J. Nie</i>	
Neighborhood Pyramid Preserving Hashing	1507
. <i>M. Wang, W. Zhou, Q. Tian, and H. Li</i>	
Learning Scene Attribute for Scene Recognition	1519
. <i>H. Zeng, X. Song, G. Chen, and S. Jiang</i>	
The Importance of Context When Recommending TV Content: Dataset and Algorithms	
. <i>M. S. Kristoffersen, S. E. Shepstone, and Z.-H. Tan</i>	1531
Unsupervised Variational Video Hashing With 1D-CNN-LSTM Networks	1542
. <i>S. Li, Z. Chen, X. Li, J. Lu, and J. Zhou</i>	
<i>Social and Web Multimedia</i>	
Low-Rank Regularized Multi-Representation Learning for Fashion Compatibility Prediction	
. <i>P. Jing, S. Ye, L. Nie, J. Liu, and Y. Su</i>	1555
<i>Cyber-Physical Multimedia Systems</i>	
Sensor-Augmented Neural Adaptive Bitrate Video Streaming on UAVs	
. <i>X. Xiao, W. Wang, T. Chen, Y. Cao, T. Jiang, and Q. Zhang</i>	1567
<i>Deep Learning for Multimedia Analysis</i>	
Coarse-to-Fine Localization of Temporal Action Proposals	1577
. <i>F. Long, T. Yao, Z. Qiu, X. Tian, T. Mei, and J. Luo</i>	
Uni-and-Bi-Directional Video Prediction via Learning Object-Centric Transformation	1591
. <i>X. Chen and W. Wang</i>	
Concentrated Local Part Discovery With Fine-Grained Part Representation for Person Re-Identification	
. <i>C. Wan, Y. Wu, X. Tian, J. Huang, and X.-S. Hua</i>	1605
<i>Deep Learning for Multimedia Processing</i>	
Asymmetric Joint GANs for Normalizing Face Illumination From a Single Image	
. <i>X. Han, H. Yang, G. Xing, and Y. Liu</i>	1619
Visual-Textual Emotion Analysis With Deep Coupled Video and Danmu Neural Networks	
. <i>C. Li, J. Wang, H. Wang, M. Zhao, W. Li, and X. Deng</i>	1634
<i>Multimedia for Social Activities in Health, Environment, and Urban Living</i>	
Hierarchical Attention Network for Visually-Aware Food Recommendation	
. <i>X. Gao, F. Feng, X. He, H. Huang, X. Guan, C. Feng, Z. Ming, and T.-S. Chua</i>	1647
