

# 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/>

JULY 2019

VOLUME 21

NUMBER 7

ITMUF8

(ISSN 1520-9210)

---

REGULAR PAPERS

*Audio/Speech/Language Analysis and Synthesis*

BLTRCNN-Based 3-D Articulatory Movement Prediction: Learning Articulatory Synchronicity From Both Text and Audio Inputs..... *L. Yu, J. Yu, and Q. Ling* 1621

*Media Conversion and Transcoding*

Fast H.264 to HEVC Transcoding: A Deep Learning Method ..... *J. Xu, M. Xu, Y. Wei, Z. Wang, and Z. Guan* 1633

*Image/Video/Graphics Analysis and Synthesis*

Saliency Detection via Multi-Scale Global Cues ..... *X. Lin, Z.-J. Wang, L. Ma, and X. Wu* 1646

HSCS: Hierarchical Sparsity Based Co-saliency Detection for RGBD Images.....  
..... *R. Cong, J. Lei, H. Fu, Q. Huang, X. Cao, and N. Ling* 1660

Unsupervised Universal Attribute Modeling for Action Recognition ..... *D. Roy, K. S. R. Murty, and C. K. Mohan* 1672

High-Quality Image Captioning With Fine-Grained and Semantic-Guided Visual Attention .....  
..... *Z. Zhang, Q. Wu, Y. Wang, and F. Chen* 1681

Dynamic Texture Classification Using Unsupervised 3D Filter Learning and Local Binary Encoding .....  
..... *X. Zhao, Y. Lin, L. Liu, J. Heikkilä, and W. Zheng* 1694

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

(Contents Continued on Back Cover)



(Contents Continued from Front Cover)

---

*Multimodal Perception, Integration, and Multisensory Fusion*

Learning a Joint Affinity Graph for Multiview Subspace Clustering . . . . . C. Tang, X. Zhu, X. Liu, M. Li, P. Wang, C. Zhang, and L. Wang 1724

*Subjective and Objective Quality Assessment and User Experience*

Blind Quality Assessment of 3-D Synthesized Views Based on Hybrid Feature Classes . . . . . F. Rodrigues, J. Ascenso, A. Rodrigues, and M. P. Queluz 1737

Predicting Stereoscopic Image Quality via Stacked Auto-Encoders Based on Stereopsis Formation . . . . . J. Yang, K. Sim, W. Lu, and B. Jiang 1750

*Big Data Support for Multimedia*

Video Big Data Retrieval Over Media Cloud: A Context-Aware Online Learning Approach . . . . . Y. Feng, P. Zhou, J. Xu, S. Ji, and D. Wu 1762

*Healthcare*

EVM-CNN: Real-Time Contactless Heart Rate Estimation From Facial Video . . . . . Y. Qiu, Y. Liu, J. Arteaga-Falconi, H. Dong, and A. El Saddik 1778

*Wireless and Mobile Multimedia*

Cache Less for More: Exploiting Cooperative Video Caching and Delivery in D2D Communications . . . . . D. Wu, Q. Liu, H. Wang, Q. Yang, and R. Wang 1788

*Deep Learning for Multimedia Analysis*

Structure-Constrained Motion Sequence Generation . . . . . Y. Yan, B. Ni, W. Zhang, J. Xu, and X. Yang 1799

Long Activity Video Understanding Using Functional Object-Oriented Network . . . . A. B. Jelodar, D. Paulius, and Y. Sun 1813

Fine-Grained Land Use Classification at the City Scale Using Ground-Level Images . . . . Y. Zhu, X. Deng, and S. Newsam 1825

*Deep Learning for Multimedia Processing*

Learning Deep Conditional Neural Network for Image Segmentation . . . . . Q. Wang, C. Yuan, and Y. Liu 1839

*Multimedia Sentiment Analysis and Synthesis; Affective Media Processing*

Inferring Emotions From Large-Scale Internet Voice Data . . . . . J. Jia, S. Zhou, Y. Yin, B. Wu, W. Chen, F. Meng, and Y. Wang 1853

*Multimedia for Personal Applications (Mobile, Wearables, Interactive)*

A Hierarchical Approach for Associating Body-Worn Sensors to Video Regions in Crowded Mingling Scenarios . . . . . L. Cabrera-Quiros and H. Hung 1867

*Speech, Language, and Audio in Video Analysis; Music in Multimedia*

A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training . . . . . R. Cui, H. Liu, and C. Zhang 1880

---

Information for Authors . . . . . <https://signalprocessingsociety.org/publications-resources/information-authors>

---