

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/>

SEPTEMBER 2020

VOLUME 22

NUMBER 9

ITMUF8

(ISSN 1520-9210)

REGULAR PAPERS

Compression and Coding

Intra Coding Strategy for Video Error Resiliency: Behavioral Analysis
..... *M. Kazemi, M. Ghanbari, and S. Shirmohammadi* 2193

Image/Video/Graphics Analysis and Synthesis

Joint Learning in the Spatio-Temporal and Frequency Domains for Skeleton-Based Action Recognition
..... *G. Hu, B. Cui, and S. Yu* 2207

Moving Cast Shadows Segmentation Using Illumination Invariant Feature *B. Wang, Y. Zhao, and C. L. Philip Chen* 2221

Learning Discriminative and Generative Shape Embeddings for Three-Dimensional Shape Retrieval
..... *C. Xu, B. Leng, B. Chen, C. Zhang, and X. Zhou* 2234

Guide to Match: Multi-Layer Feature Matching With a Hybrid Gaussian Mixture Model
..... *K. Sun, W. Tao, and Y. Qian* 2246

3-D Processing and Presentation

Transfer Function-Guided Saliency-Aware Compression for Transmitting Volumetric Data
..... *J. H. Park, I. Gutenko, and A. E. Kaufman* 2262

A Unified Deep Metric Representation for Mesh Saliency Detection and Non-Rigid Shape Matching
..... *S. Hu, H. P. H. Shum, N. Aslam, F. W. B. Li, and X. Liang* 2278

Video Surveillance and Semantic Analysis

Convolutional Networks With Channel and STIPs Attention Model for Action Recognition in Videos
..... *H. Wu, X. Ma, and Y. Li* 2293

Multimodal Human-Machine Interfaces and Interaction

Visual Relationship Embedding Network for Image Paragraph Generation *W. Che, X. Fan, R. Xiong, and D. Zhao* 2307

(Contents Continued on Back Cover)



<i>Subjective and Objective Quality Assessment and User Experience</i>	
QoE-Aware Multi-Source Video Streaming in Content Centric Networks	<i>M. N. Sadat, R. Dai, L. Kong, and J. Zhu</i> 2321
<i>Virtual and Augmented Reality</i>	
The Prediction of Saliency Map for Head and Eye Movements in 360 Degree Images	<i>Y. Zhu, G. Zhai, X. Min, and J. Zhou</i> 2331
<i>Multimedia Search and Retrieval</i>	
Semi-Supervised Cross-Modal Retrieval With Label Prediction	<i>D. Mandal, P. Rao, and S. Biswas</i> 2345
Rich Features Embedding for Cross-Modal Retrieval: A Simple Baseline	<i>X. Fu, Y. Zhao, Y. Wei, Y. Zhao, and S. Wei</i> 2354
<i>Immersive Communications and Networking</i>	
A Fast FoV-Switching DASH System Based on Tiling Mechanism for Practical Omnidirectional Video Services	<i>J. Song, F. Yang, W. Zhang, W. Zou, Y. Fan, and P. Di</i> 2366
Tile-Based Joint Caching and Delivery of 360° Videos in Heterogeneous Networks	<i>P. Maniotis, E. Bourtsoulatze, and N. Thomos</i> 2382
<i>Multimedia Streaming and Transport</i>	
An Efficient NVoD Scheme Using Implicit Error Correction and Subchannels for Wireless Networks	<i>R. Asorey-Cacheda, A.-J. Garcia-Sanchez, and J. Garcia-Haro</i> 2396
A Distance-Driven Alliance for a P2P Live Video System	<i>J. Zhang, Y. Zhang, and M. Shen</i> 2409
<i>Deep Learning for Multimedia Analysis</i>	
Weighted and Class-Specific Maximum Mean Discrepancy for Unsupervised Domain Adaptation	<i>H. Yan, Z. Li, Q. Wang, P. Li, Y. Xu, and W. Zuo</i> 2420
An Attentive Sequence to Sequence Translator for Localizing Video Clips by Natural Language	<i>K. Ning, M. Cai, D. Xie, and F. Wu</i> 2434
Leveraging Virtual and Real Person for Unsupervised Person Re-Identification	<i>F. Yang, Z. Zhong, Z. Luo, S. Lian, and S. Li</i> 2444
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>	
Affective Video Content Analysis With Adaptive Fusion Recurrent Network	<i>Y. Yi, H. Wang, and Q. Li</i> 2454
<i>User-Centric Social Multimedia Computing and Location Based Multimedia Service</i>	
Massive-Scale Genre Communities Learning Using a Noise-Tolerant Deep Architecture	<i>L. Zhang, X. Ju, Y. Yao, and Z. Liu</i> 2467
