

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/>

2021

VOLUME 23

ITMUF8

(ISSN 1520-9210)

REGULAR PAPERS

Audio/Speech/Language Analysis and Synthesis

End-to-End Audiovisual Speech Recognition System With Multitask Learning *F. Tao and C. Busso* 1

Compression and Coding

Soft Video Multicasting Using Adaptive Compressed Sensing *H. Hadizadeh and I. V. Bajić* 12

BVI-SynTex: A Synthetic Video Texture Dataset for Video Compression and Quality Assessment
..... *A. V. Katsenou, G. Dimitrov, D. Ma, and D. R. Bull* 26

Fast Non-Local Adaptive In-Loop Filter Optimization on GPU
..... *C. Jia, F. Luo, X. Zhang, S. Wang, S. Wang, and S. Ma* 39

Watermarking, Encryption, and Data Hiding

High-Fidelity Reversible Image Watermarking Based on Effective Prediction Error-Pairs Modification
..... *W. He, Z. Cai, and Y. Wang* 52

Image/Video/Graphics Analysis and Synthesis

A Multi-Stream Graph Convolutional Networks-Hidden Conditional Random Field Model for Skeleton-Based Action
Recognition *K. Liu, L. Gao, N. M. Khan, L. Qi, and L. Guan* 64

Constant Size Point Cloud Clustering: A Compact, Non-Overlapping Solution
..... *A. F. R. Guarda, N. M. M. Rodrigues, and F. Pereira* 77

Integrating Part of Speech Guidance for Image Captioning *J. Zhang, K. Mei, Y. Zheng, and J. Fan* 92

A Bottom-Up and Top-Down Integration Framework for Online Object Tracking *M. Li, L. Peng, T. Wu, and Z. Peng* 105

Siamese Tracking Network With Informative Enhanced Loss *S. Tian, X. Liu, M. Liu, S. Li, and B. Yin* 120

Pose-Normalized and Appearance-Preserved Street-to-Shop Clothing Image Generation and Feature Learning
..... *H. Zhan, C. Yi, B. Shi, J. Lin, L.-Y. Duan, and A. C. Kot* 133

Adversarial Disentanglement Spectrum Variations and Cross-Modality Attention Networks for NIR-VIS Face Recognition
..... *W. Hu and H. Hu* 145

Pose-Guided Tracking-by-Detection: Robust Multi-Person Pose Tracking
..... *Q. Bao, W. Liu, Y. Cheng, B. Zhou, and T. Mei* 161

(Contents Continued on Page i)



Learning Representations for High-Dynamic-Range Image Color Transfer in a Self-Supervised Way	176
. <i>Y. Huang, S. Qiu, C. Wang, and C. Li</i>	
Enhancing Underexposed Photos Using Perceptually Bidirectional Similarity	189
. <i>Q. Zhang, Y. Nie, L. Zhu, C. Xiao, and W.-S. Zheng</i>	
Spatial-Temporal Cascade Autoencoder for Video Anomaly Detection in Crowded Scenes	203
. <i>N. Li, F. Chang, and C. Liu</i>	
<i>Sparse Multimedia Signal Processing</i>	
Learning Adaptive Neighborhood Graph on Grassmann Manifolds for Video/Image-Set Subspace Clustering	216
. <i>B. Wang, Y. Hu, J. Gao, Y. Sun, F. Ju, and B. Yin</i>	
Graph Embedding Multi-Kernel Metric Learning for Image Set Classification With Grassmannian Manifold-Valued Features	228
. <i>R. Wang, X.-J. Wu, and J. Kittler</i>	
<i>Large-Scale Multimedia Systems and Benchmarking</i>	
Interactive Video Retrieval in the Age of Deep Learning – Detailed Evaluation of VBS 2019	243
. <i>L. Rossetto, R. Gasser, J. Lokoč, W. Bailer, K. Schoeffmann, B. Muenzer, T. Souček, P. A. Nguyen, P. Bolettieri, A. Leibetseder, and S. Vrochidis</i>	
<i>3-D Processing and Presentation</i>	
Viewpoint Recommendation Based on Object-Oriented 3D Scene Reconstruction	257
. <i>K. Li, Y. Wu, Y. Xue, and X. Qian</i>	
<i>Video Surveillance and Semantic Analysis</i>	
Hierarchical Reasoning Network for Pedestrian Attribute Recognition	268
. <i>H. An, H.-M. Hu, Y. Guo, Q. Zhou, and B. Li</i>	
Person Re-Identification in Aerial Imagery	281
. <i>S. Zhang, Q. Zhang, Y. Yang, X. Wei, P. Wang, B. Jiao, and Y. Zhang</i>	
<i>Multimodal Perception, Integration, and Multisensory Fusion</i>	
Re-Synchronization Using the Hand Preceding Model for Multi-Modal Fusion in Automatic Continuous Cued Speech Recognition	292
. <i>L. Liu, G. Feng, D. Beautemps, and X.-P. Zhang</i>	
Spectrum Characteristics Preserved Visible and Near-Infrared Image Fusion Algorithm	306
. <i>Z. Li, H.-M. Hu, W. Zhang, S. Pu, and B. Li</i>	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Quality Index for View Synthesis by Measuring Instance Degradation and Global Appearance	320
. <i>L. Li, Y. Zhou, J. Wu, F. Li, and G. Shi</i>	
A Physiology-Based QoE Comparison of Interactive Augmented Reality, Virtual Reality and Tablet-Based Applications	333
. <i>C. Keighrey, R. Flynn, S. Murray, and N. Murray</i>	
<i>Multimedia Search and Retrieval</i>	
Fast Nearest Subspace Search via Random Angular Hashing	342
. <i>Y. Xu, X. Liu, B. Wang, R. Tao, K. Xia, and X. Cao</i>	
Boosting Temporal Binary Coding for Large-Scale Video Search	353
. <i>Y. Wu, X. Liu, H. Qin, K. Xia, S. Hu, Y. Ma, and M. Wang</i>	
Dress With Style: Learning Style From Joint Deep Embedding of Clothing Styles and Body Shapes	365
. <i>S. C. Hidayati, T. W. Goh, J.-S. G. Chan, C.-C. Hsu, J. See, L.-K. Wong, K.-L. Hua, Y. Tsao, and W.-H. Cheng</i>	
<i>Social and Web Multimedia</i>	
LAST: Location-Appearance-Semantic-Temporal Clustering Based POI Summarization	378
. <i>X. Qian, Y. Wu, M. Li, Y. Ren, S. Jiang, and Z. Li</i>	
<i>Deep Learning for Multimedia Analysis</i>	
SPA-GAN: Spatial Attention GAN for Image-to-Image Translation	391
. <i>H. Emami, M. M. Aliabadi, M. Dong, and R. B. Chinnam</i>	
Learning Localized Representations of Point Clouds With Graph-Convolutional Generative Adversarial Networks	402
. <i>D. Valsesia, G. Fracastoro, and E. Magli</i>	
3-D Human Behavior Understanding Using Generalized TS-LSTM Networks	415
. <i>I. Lee, D. Kim, and S. Lee</i>	
Recurrent Generative Adversarial Network for Face Completion	429
. <i>Q. Wang, H. Fan, G. Sun, W. Ren, and Y. Tang</i>	
Density-Aware Multi-Task Learning for Crowd Counting	443
. <i>X. Jiang, L. Zhang, T. Zhang, P. Lv, B. Zhou, Y. Pang, M. Xu, and C. Xu</i>	
<i>Deep Learning for Multimedia Processing</i>	
OPMP: An Omnidirectional Pyramid Mask Proposal Network for Arbitrary-Shape Scene Text Detection	454
. <i>S. Zhang, Y. Liu, L. Jin, Z. Wei, and C. Shen</i>	

Learning Face Image Super-Resolution Through Facial Semantic Attribute Transformation and Self-Attentive Structure Enhancement	<i>M. Li, Z. Zhang, J. Yu, and C. W. Chen</i>	468
<i>Multimedia for Immersive Search Space and Personalized Recommendations</i>		
Learning and Fusing Multiple User Interest Representations for Micro-Video and Movie Recommendations	<i>X. Chen, D. Liu, Z. Xiong, and Z.-J. Zha</i>	484
<i>Speech, Language, and Audio in Video Analysis; Music in Multimedia</i>		
DeepDance: Music-to-Dance Motion Choreography With Adversarial Learning	<i>G. Sun, Y. Wong, Z. Cheng, M. S. Kankanhalli, W. Geng, and X. Li</i>	497
<i>User-Centric Social Multimedia Computing and Location Based Multimedia Service</i>		
Metadata Connector: Exploiting Hashtag and Tag for Cross-OSN Event Search	<i>Y. Gao, J. Sang, C. Fu, Z. Wang, T. Ren, and C. Xu</i>	510
<i>Deep Learning for Multimedia Analysis</i>		
A Novel Perspective to Zero-Shot Learning: Towards an Alignment of Manifold Structures via Semantic Feature Expansion	<i>J. Guo and S. Guo</i>	524
<i>Multimedia Search and Retrieval</i>		
IPTV Channel Zapping Recommendation With Attention Mechanism	<i>G. Li, L. Qiu, C. Yu, H. Cao, Y. Liu, and C. Yang</i>	538
<i>Deep Learning for Multimedia Processing</i>		
Mask Cross-Modal Hashing Networks	<i>Q. Lin, W. Cao, Z. He, and Z. He</i>	550
<i>Multimedia Storytelling and Cross-Modal Translations Between Multimedia Contents</i>		
Augmented Adversarial Training for Cross-Modal Retrieval	<i>Y. Wu, S. Wang, G. Song, and Q. Huang</i>	559
<i>Deep Learning for Multimedia Processing</i>		
Driver Yawning Detection Based on Subtle Facial Action Recognition	<i>H. Yang, L. Liu, W. Min, X. Yang, and X. Xiong</i>	572
<i>Cloud Media Platforms and Innovative Applications and Cloud Gaming</i>		
Learned Resolution Scaling Powered Gaming-as-a-Service at Scale	<i>H. Chen, M. Lu, Z. Ma, X. Zhang, Y. Xu, Q. Shen, and W. Zhang</i>	584
<i>Video Surveillance and Semantic Analysis</i>		
Progressive Unsupervised Person Re-Identification by Tracklet Association With Spatio-Temporal Regularization	<i>Q. Xie, W. Zhou, G.-J. Qi, Q. Tian, and H. Li</i>	597
<i>Subjective and Objective Quality Assessment and User Experience</i>		
Beyond Vision: A Multimodal Recurrent Attention Convolutional Neural Network for Unified Image Aesthetic Prediction Tasks	<i>X. Zhang, X. Gao, W. Lu, L. He, and J. Li</i>	611
<i>Deep Learning for Multimedia Processing</i>		
BR ² Net: Defocus Blur Detection Via a Bidirectional Channel Attention Residual Refining Network	<i>C. Tang, X. Liu, S. An, and P. Wang</i>	624
<i>Watermarking, Encryption, and Data Hiding</i>		
A Recursive Reversible Data Hiding in Encrypted Images Method With a Very High Payload	<i>P. Puteaux and W. Puech</i>	636
<i>Real-Time Communications</i>		
TCLiVi: Transmission Control in Live Video Streaming Based on Deep Reinforcement Learning	<i>L. Cui, D. Su, S. Yang, Z. Wang, and Z. Ming</i>	651
<i>Image/Video/Graphics Analysis and Synthesis</i>		
Utilizing Two-Phase Processing With FBLS for Single Image Deraining	<i>X. Lin, L. Ma, B. Sheng, Z.-J. Wang, and W. Chen</i>	664
<i>Multimedia Search and Retrieval</i>		
Benchmarking Image Retrieval Diversification Techniques for Social Media	<i>B. Ionescu, M. Rohm, B. Boteanu, A. L. Gînscă, M. Lupu, and H. Müller</i>	677

<i>Subjective and Objective Quality Assessment and User Experience</i>	
Exploiting Local Degradation Characteristics and Global Statistical Properties for Blind Quality Assessment of Tone-Mapped HDR Images	X. Wang, Q. Jiang, F. Shao, K. Gu, G. Zhai, and X. Yang 692
<i>Audio/Speech/Language Analysis and Synthesis</i>	
A Hierarchical Visual Feature-Based Approach For Image Sonification	O. K. Toffa and M. Mignotte 706
<i>Virtual and Augmented Reality</i>	
Predictive Adaptive Streaming to Enable Mobile 360-Degree and VR Experiences	X. Hou, S. Dey, J. Zhang, and M. Budagavi 716
<i>Sparse Multimedia Signal Processing</i>	
Patch Based Video Summarization With Block Sparse Representation	S. Mei, M. Ma, S. Wan, J. Hou, Z. Wang, and D. D. Feng 732
<i>Virtual and Augmented Reality</i>	
Viewport-Dependent Saliency Prediction in 360° Video	M. Qiao, M. Xu, Z. Wang, and A. Borji 748
<i>Deep Learning for Multimedia Processing</i>	
Learning Crisp Boundaries Using Deep Refinement Network and Adaptive Weighting Loss	Y.-J. Cao, C. Lin, and Y.-J. Li 761
Frequency-Dependent Depth Map Enhancement via Iterative Depth-Guided Affine Transformation and Intensity-Guided Refinement	Y. Zuo, Y. Fang, P. An, X. Shang, and J. Yang 772
<i>Multimedia Interfaces and Interaction</i>	
An Attention-Based Unsupervised Adversarial Model for Movie Review Spam Detection	Y. Gao, M. Gong, Y. Xie, and A. K. Qin 784
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Panoramic Video Quality Assessment Based on Non-Local Spherical CNN	J. Yang, T. Liu, B. Jiang, W. Lu, and Q. Meng 797
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Intermittent Contextual Learning for Keyfilter-Aware UAV Object Tracking Using Deep Convolutional Feature	Y. Li, C. Fu, Z. Huang, Y. Zhang, and J. Pan 810
<i>System Design and Optimization</i>	
Energy Optimization and QoE Satisfaction for Wireless Visual Sensor Networks in Multi Target Tracking Scenario	R. Ghazalian, A. Aghagolzadeh, and S. M. H. Andargoli 823
<i>Multimedia Storytelling and Cross-Modal Translations Between Multimedia Contents</i>	
CaptionNet: A Tailor-made Recurrent Neural Network for Generating Image Descriptions	L. Yang, H. Wang, P. Tang, and Q. Li 835
<i>Deep Learning for Multimedia Analysis</i>	
Single Shot Video Object Detector	J. Deng, Y. Pan, T. Yao, W. Zhou, H. Li, and T. Mei 846
<i>3-D Video Signal Processing</i>	
CAT: Corner Aided Tracking With Deep Regression Network	S. Zhang, X. Zhao, and L. Fang 859
<i>Deep Learning for Multimedia Processing</i>	
Progressive Learning of Low-Precision Networks for Image Classification	Z. Zhou, W. Zhou, X. Lv, X. Huang, X. Wang, and H. Li 871
<i>Video Surveillance and Semantic Analysis</i>	
Hierarchical Soft Quantization for Skeleton-Based Human Action Recognition	J. Yang, W. Liu, J. Yuan, and T. Mei 883
<i>Multimedia Search and Retrieval</i>	
A Mutually Attentive Co-Training Framework for Semi-Supervised Recognition	S. Min, X. Chen, H. Xie, Z.-J. Zha, and Y. Zhang 899

<i>Multimedia Streaming and Transport</i>	
Modeling QoE for Buffered Video Streaming in Interference-Limited Cellular Networks	911
..... P. Schulz, H. Klessig, M. Simsek, and G. Fettweis	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
A Coarse-to-Fine Facial Landmark Detection Method Based on Self-attention Mechanism	926
..... P. Gao, K. Lu, J. Xue, L. Shao, and J. Lyu	
<i>Multimedia Search and Retrieval</i>	
Deep Loss Driven Multi-Scale Hashing Based on Pyramid Connected Network	939
..... L. Gu, J. Liu, X. Liu, and J. Sun	
<i>HDR and Ultra HD Multimedia Processing and Communications</i>	
Blind Quality Assessment for Tone-Mapped Images by Analysis of Gradient and Chromatic Statistics	955
..... Y. Fang, J. Yan, R. Du, Y. Zuo, W. Wen, Y. Zeng, and L. Li	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Attentive Cross-Modal Fusion Network for RGB-D Saliency Detection	967
..... D. Liu, K. Zhang, and Z. Chen	
<i>Multimedia Content Delivery Networks</i>	
QoS-Aware Multicast for Scalable Video Streaming in Software-Defined Networks	982
..... H. Zhong, F. Wu, Y. Xu, and J. Cui	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Query Reconstruction Network for Referring Expression Image Segmentation ...	995
..... H. Shi, H. Li, Q. Wu, and K. N. Ngan	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Semi-Reference Sonar Image Quality Assessment Based on Task and Visual Perception	1008
..... W. Chen, K. Gu, T. Zhao, G. Jiang, and P. L. Callet	
<i>Deep Learning for Multimedia Processing</i>	
3D Pose Estimation Based on Reinforce Learning for 2D Image-Based 3D Model Retrieval	1021
..... W.-Z. Nie, W.-W. Jia, W.-H. Li, A.-A. Liu, and S.-C. Zhao	
<i>Deep Learning for Multimedia Analysis</i>	
SAL: Selection and Attention Losses for Weakly Supervised Semantic Segmentation	1035
..... L. Zhou, C. Gong, Z. Liu, and K. Fu	
<i>Display Technology for Multimedia</i>	
Deep Battery Saver: End-to-End Learning for Power Constrained Contrast Enhancement	1049
..... J.-L. Yin, B.-H. Chen, Y.-T. Peng, and C.-C. Tsai	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
DENet: A Universal Network for Counting Crowd With Varying Densities and Scales	1060
..... L. Liu, J. Jiang, W. Jia, S. Amirgholipour, Y. Wang, M. Zeibots, and X. He	
<i>Media Conversion and Transcoding</i>	
A New Image Compression Algorithm Based on Non-Uniform Partition and U-System	1069
..... Y. Zhang, Z. Cai, and G. Xiong	
<i>Deep Learning for Multimedia Analysis</i>	
Adversarial Learning for Personalized Tag Recommendation	1083
..... E. Quintanilla, Y. Rawat, A. Sakryukin, M. Shah, and M. Kankanhalli	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Less is (Just as Good as) More - an Investigation of Odor Intensity and Hedonic Valence in Mulsemmedia QoE using Heart Rate and Eye Tracking	1095
..... G. Mesfin, E. B. Saleme, O. A. Ademoye, E. Kani-Zabihi, C. A.S. Santos, and G. Ghinea	
<i>Compression and Coding</i>	
Rate Control Method Based on Deep Reinforcement Learning for Dynamic Video Sequences in HEVC	1106
..... M. Zhou, X. Wei, S. Kwong, W. Jia, and B. Fang	

<i>Low-Power Digital and Analog Circuits for Multimedia</i>	
A 460 GOPS/W Improved Mnemonic Descent Method-Based Hardwired Accelerator for Face Alignment	1122
. <i>H. Mo, L. Liu, W. Zhu, Q. Li, S. Yin, and S. Wei</i>	
<i>Healthcare</i>	
An Artificial Intelligence-Based System to Assess Nutrient Intake for Hospitalised Patients	1136
. <i>Y. Lu, T. Stathopoulou, M. F. Vasiloglou, S. Christodoulidis, Z. Stanga, and S. Mougiakakou</i>	
<i>Sparse Multimedia Signal Processing</i>	
Probabilistic Undirected Graph Based Denoising Method for Dynamic Vision Sensor	1148
. <i>J. Wu, C. Ma, L. Li, W. Dong, and G. Shi</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
3D Face Reconstruction From A Single Image Assisted by 2D Face Images in the Wild	1160
. <i>X. Tu, J. Zhao, M. Xie, Z. Jiang, A. Balamurugan, Y. Luo, Y. Zhao, L. He, Z. Ma, and J. Feng</i>	
<i>3-D Processing and Presentation</i>	
Measuring Coarse-to-Fine Texture and Geometric Distortions for Quality Assessment of DIBR-Synthesized Images	1173
. <i>X. Wang, F. Shao, Q. Jiang, X. Meng, and Y.-S. Ho</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Fine-Grained Visual Categorization by Localizing Object Parts With Single Image	1187
. <i>X. Zheng, L. Qi, Y. Ren, and X. Lu</i>	
<i>Multimedia Search and Retrieval</i>	
Attribute-Guided Feature Learning for Few-Shot Image Recognition	1200
. <i>Y. Zhu, W. Min, and S. Jiang</i>	
Complementary Incremental Hashing With Query-Adaptive Re-Ranking for Image Retrieval	1210
. <i>X. Tian, W. W. Y. Ng, H. Wang, and S. Kwong</i>	
<i>Sparse Multimedia Signal Processing</i>	
Factorized Tensor Dictionary Learning for Visual Tensor Data Completion	1225
. <i>R. Xu, Y. Xu, and Y. Quan</i>	
<i>Multimedia Search and Retrieval</i>	
Progressive Bilateral-Context Driven Model for Post-Processing Person Re-Identification	1239
. <i>M. Cao, C. Chen, H. Dou, X. Hu, S. Peng, and A. Kuijper</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Dual Neural Networks Coupling Data Regression With Explicit Priors for Monocular 3D Face Reconstruction	1252
. <i>X. Fan, S. Cheng, K. Huyan, M. Hou, R. Liu, and Z. Luo</i>	
<i>Deep Learning for Multimedia Processing</i>	
Self-Adaptive Neural Module Transformer for Visual Question Answering	1264
. <i>H. Zhong, J. Chen, C. Shen, H. Zhang, J. Huang, and X.-S. Hua</i>	
<i>Deep Learning for Multimedia Analysis</i>	
Deep Collaborative Discrete Hashing With Semantic-Invariant Structure Construction	1274
. <i>Z. Wang, Z. Zhang, Y. Luo, Z. Huang, and H. T. Shen</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Object Cosegmentation in Noisy Videos With Multilevel Hypergraph	1287
. <i>L. Wang, X. Lv, Q. Zhang, Z. Niu, N. Zheng, and G. Hua</i>	
A Novel Image Representation Method Under a Non-Standard Positional Numeral System	1301
. <i>T. Lan and Z. Cai</i>	
<i>Deep Learning for Multimedia Processing</i>	
R-Net: A Relationship Network for Efficient and Accurate Scene Text Detection	1316
. <i>Y. Wang, H. Xie, Z. Zha, Y. Tian, Z. Fu, and Y. Zhang</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Hybrid Refinement-Correction Heatmaps for Human Pose Estimation	1330
. <i>A. Kamel, B. Sheng, P. Li, J. Kim, and D. D. Feng</i>	
<i>Video Surveillance and Semantic Analysis</i>	
cmSalGAN: RGB-D Salient Object Detection With Cross-View Generative Adversarial Networks	1343
. <i>B. Jiang, Z. Zhou, X. Wang, J. Tang, and B. Luo</i>	

<i>Deep Learning for Multimedia Analysis</i>	
Snowball: Iterative Model Evolution and Confident Sample Discovery for Semi-Supervised Learning on Very Small Labeled Datasets	Y. Li, Z. Zhao, H. Sun, Y. Cen, and Z. He 1354
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Prominent Local Representation for Dynamic Textures Based on High-Order Gaussian-Gradients	T. T. Nguyen, T. P. Nguyen, and F. Bouchara 1367
AttentionFGAN: Infrared and Visible Image Fusion Using Attention-Based Generative Adversarial Networks	J. Li, H. Huo, C. Li, R. Wang, and Q. Feng 1383
<i>Deep Learning for Multimedia Processing</i>	
Stacked U-Shape Network With Channel-Wise Attention for Salient Object Detection	J. Li, Z. Pan, Q. Liu, and Z. Wang 1397
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Image-Only Real-Time Incremental UAV Image Mosaic for Multi-Strip Flight	F. Zhang, T. Yang, L. F. Liu, B. Liang, Y. Bai, and J. Li 1410
Weighted Adaptive Image Super-Resolution Scheme Based on Local Fractal Feature and Image Roughness	X. Yao, Q. Wu, P. Zhang, and F. Bao 1426
<i>Deep Learning for Multimedia Processing</i>	
Salient Object Detection by Fusing Local and Global Contexts	Q. Ren, S. Lu, J. Zhang, and R. Hu 1442
<i>Multimodal Human-Machine Interfaces and Interaction</i>	
A Brain-Media Deep Framework Towards Seeing Imaginations Inside Brains	J. Jiang, A. Fares, and S.-H. Zhong 1454
<i>Watermarking, Encryption, and Data Hiding</i>	
High Capacity Reversible Data Hiding in Encrypted Image Based on Intra-Block Lossless Compression	Y. Wang, Z. Cai, and W. He 1466
<i>Deep Learning for Multimedia Analysis</i>	
Deep Multi-Patch Matching Network for Visible Thermal Person Re-Identification	P. Wang, Z. Zhao, F. Su, Y. Zhao, H. Wang, L. Yang, and Y. Li 1474
<i>Deep Learning for Multimedia Processing</i>	
Coarse-to-Fine CNN for Image Super-Resolution	C. Tian, Y. Xu, W. Zuo, B. Zhang, L. Fei, and C.-W. Lin 1489
<i>3-D Video Signal Processing</i>	
Transferable Knowledge-Based Multi-Granularity Fusion Network for Weakly Supervised Temporal Action Detection	H. Su, X. Zhao, T. Lin, S. Liu, and Z. Hu 1503
<i>Compression and Coding</i>	
Perceptual Image Hashing With Texture and Invariant Vector Distance for Copy Detection	Z. Huang and S. Liu 1516
<i>Media Cloud Computing and Communications</i>	
QoE-driven HAS Live Video Channel Placement in the Media Cloud	J. Liu, W. Zhang, S. Huang, H. Du, and Q. Zheng 1530
<i>Multimedia for Augmented Experience in Real and Virtual Life</i>	
Building High-Fidelity Human Body Models From User-Generated Data	Z. Xu, W. Chang, Y. Zhu, L. Dong, H. Zhou, and Q. Zhang 1542
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Blind Image Quality Assessment Based on Multi-scale KLT	C. Yang, X. Zhang, P. An, L. Shen, and C.-C. J. Kuo 1557
<i>Multimedia for Personal Applications (Mobile, Wearables, Interactive)</i>	
V-Eye: A Vision-Based Navigation System for the Visually Impaired	P.-J. Duh, Y.-C. Sung, L.-Y. F. Chiang, Y.-J. Chang, and K.-W. Chen 1567
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Orthogonalization-Guided Feature Fusion Network for Multimodal 2D+3D Facial Expression Recognition	S. Lin, M. Bai, F. Liu, L. Shen, and Y. Zhou 1581

<i>Sparse Multimedia Signal Processing</i>	
Image Quality Assessment Using Kernel Sparse Coding	Z. Zhou, J. Li, Y. Quan, and R. Xu 1592
<i>Multimedia Search and Retrieval</i>	
Tag Propagation and Cost-Sensitive Learning for Music Auto-Tagging	Y.-H. Lin and H. H. Chen 1605
<i>3-D Video Signal Processing</i>	
SparseFusion: Dynamic Human Avatar Modeling From Sparse RGBD Images	X. Zuo, S. Wang, J. Zheng, W. Yu, M. Gong, R. Yang, and L. Cheng 1617
<i>Image/Video/Graphics Analysis and Synthesis</i>	
A Novel Depth and Color Feature Fusion Framework for 6D Object Pose Estimation	G. Zhou, Y. Yan, D. Wang, and Q. Chen 1630
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>	
Adaptive Deep Metric Learning for Affective Image Retrieval and Classification	X. Yao, D. She, H. Zhang, J. Yang, M.-M. Cheng, and L. Wang 1640
<i>Immersive Communications and Networking</i>	
Semantic Example Guided Image-to-Image Translation	J. Huang, J. Liao, and S. Kwong 1654
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Low-Rank Pairwise Alignment Bilinear Network For Few-Shot Fine-Grained Image Classification	H. Huang, J. Zhang, J. Zhang, J. Xu, and Q. Wu 1666
<i>Video Surveillance and Semantic Analysis</i>	
Part-aware Progressive Unsupervised Domain Adaptation for Person Re-Identification	F. Yang, K. Yan, S. Lu, H. Jia, D. Xie, Z. Yu, X. Guo, F. Huang, and W. Gao 1681
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Joint Input and Output Space Learning for Multi-Label Image Classification	J. Xu, H. Tian, Z. Wang, Y. Wang, W. Kang, and F. Chen 1696
<i>Multimedia Search and Retrieval</i>	
Real-world Cross-modal Retrieval via Sequential Learning	G. Song and X. Tan 1708
<i>Deep Learning for Multimedia Analysis</i>	
COMO: Efficient Deep Neural Networks Expansion With CONvolutional MaxOut	B. Zhao, H. Xiong, J. Bian, Z. Guo, C.-Z. Xu, and D. Dou 1722
<i>Multimedia Storytelling and Cross-Modal Translations Between Multimedia Contents</i>	
Learning to Visualize Music Through Shot Sequence for Automatic Concert Video Mashup	W.-L. Wei, J.-C. Lin, T.-L. Liu, H.-R. Tyan, H.-M. Wang, and H.-Y. M. Liao 1731
<i>Deep Learning for Multimedia Analysis</i>	
Adversarial Multimodal Network for Movie Story Question Answering	Z. Yuan, S. Sun, L. Duan, C. Li, X. Wu, and C. Xu 1744
<i>Wireless and Mobile Multimedia</i>	
Risk Optimization for Revenue-Driven Wireless Video Broadcasting Systems: A Copula-Based Framework	W. Ji and H. V. Poor 1757
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Deep Reinforcement Polishing Network for Video Captioning	W. Xu, J. Yu, Z. Miao, L. Wan, Y. Tian, and Q. Ji 1772
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>	
LD-MAN: Layout-Driven Multimodal Attention Network for Online News Sentiment Recognition	W. Guo, Y. Zhang, X. Cai, L. Meng, J. Yang, and X. Yuan 1785
<i>Deep Learning for Multimedia Analysis</i>	
Dense Video Captioning Using Graph-Based Sentence Summarization	Z. Zhang, D. Xu, W. Ouyang, and L. Zhou 1799

<i>Virtual and Augmented Reality</i>	
A Multi-FoV Viewport-Based Visual Saliency Model Using Adaptive Weighting Losses for 360° Images	1811
. <i>F.-Y. Chao, L. Zhang, W. Hamidouche, and O. Déforges</i>	
<i>Deep Learning for Multimedia Processing</i>	
Disentangling, Embedding and Ranking Label Cues for Multi-Label Image Recognition	1827
. <i>Z.-M. Chen, Q. Cui, X.-S. Wei, X. Jin, and Y. Guo</i>	
<i>Multimodal Human-machine Interfaces and Interaction</i>	
3D Skeletal Gesture Recognition via Discriminative Coding on Time-Warping Invariant Riemannian Trajectories	1841
. <i>X. Liu and G. Zhao</i>	
<i>Multimodal Perception, Integration, and Multisensory Fusion</i>	
Shared Low-Rank Correlation Embedding for Multiple Feature Fusion	1855
. <i>Z. Wang, L. Wang, J. Wan, and H. Huang</i>	
<i>Multimedia Search and Retrieval</i>	
Online Hashing With Bit Selection for Image Retrieval	1868
. <i>Z. Weng and Y. Zhu</i>	
<i>Multimodal Perception, Integration, and Multisensory Fusion</i>	
Learning Feature Representation and Partial Correlation for Multimodal Multi-Label Data	1882
. <i>G. Song, S. Wang, Q. Huang, and Q. Tian</i>	
<i>Audio/Speech/Language Analysis and Synthesis</i>	
AutoFoley: Artificial Synthesis of Synchronized Sound Tracks for Silent Videos With Deep Learning	1895
. <i>S. Ghose and J. J. Prevost</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Motion Compensated Virtual View Synthesis Using Novel Particle Cell	1908
. <i>C. H. Cheung, L. Sheng, and K. N. Ngan</i>	
Accurate and Efficient Image Super-Resolution via Global-Local Adjusting Dense Network	1924
. <i>X. Zhang, P. Gao, S. Liu, K. Zhao, G. Li, L. Yin, and C. W. Chen</i>	
Supervised Pixel-Wise GAN for Face Super-Resolution	1938
. <i>M. Zhang and Q. Ling</i>	
<i>Watermarking, Encryption, and Data Hiding</i>	
An Automated and Robust Image Watermarking Scheme Based on Deep Neural Networks	1951
. <i>X. Zhong, P.-C. Huang, S. Mastorakis, and F. Y. Shih</i>	
<i>Deep Learning for Multimedia Analysis</i>	
M-GCN: Multi-Branch Graph Convolution Network for 2D Image-based on 3D Model Retrieval	1962
. <i>W.-Z. Nie, M.-J. Ren, A.-A. Liu, Z. Mao, and J. Nie</i>	
<i>Multimedia for Immersive Search Space and Personalized Recommendations</i>	
Neural Collaborative Preference Learning With Pairwise Comparisons	1977
. <i>Z. Li, Q. Xu, Y. Jiang, K. Ma, X. Cao, and Q. Huang</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Character Detection in Animated Movies Using Multi-Style Adaptation and Visual Attention	1990
. <i>H. Kim, E.-C. Lee, Y. Seo, D.-H. Im, and I.-K. Lee</i>	
Unsupervised Moving Object Detection in Complex Scenes Using Adversarial Regularizations	2005
. <i>M. Sultana, A. Mahmood, and S. K. Jung</i>	
<i>Multimedia for Immersive Search Space and Personalized Recommendations</i>	
Context-Dependent Propagating-Based Video Recommendation in Multimodal Heterogeneous Information Networks	2019
. <i>L. Sang, M. Xu, S. Qian, M. Martin, P. Li, and X. Wu</i>	
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>	
Weakly Supervised Emotion Intensity Prediction for Recognition of Emotions in Images	2033
. <i>H. Zhang and M. Xu</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Semantic Context Encoding for Accurate 3D Point Cloud Segmentation	2045
. <i>H. Liu, Y. Guo, Y. Ma, Y. Lei, and G. Wen</i>	
Manifold Transfer Learning via Discriminant Regression Analysis	2056
. <i>Y. Lu, W. Wang, C. Yuan, X. Li, and Z. Lai</i>	

<i>Computational Social Science and Multimedia</i>	
RealVAD: A Real-World Dataset and A Method for Voice Activity Detection by Body Motion Analysis	2071
. <i>C. Beyan, M. Shahid, and V. Murino</i>	
<i>Deep Learning for Multimedia Analysis</i>	
Understanding More About Human and Machine Attention in Deep Neural Networks	2086
. <i>Q. Lai, S. Khan, Y. Nie, H. Sun, J. Shen, and L. Shao</i>	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Subjective and Objective Quality Assessment for Stereoscopic Image Retargeting	2100
. <i>Z. Fu, F. Shao, Q. Jiang, X. Meng, and Y.-S. Ho</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Learning Deep Multi-Level Similarity for Thermal Infrared Object Tracking	2114
. <i>Q. Liu, X. Li, Z. He, N. Fan, D. Yuan, and H. Wang</i>	
<i>Multimedia Search and Retrieval</i>	
Joint Intermediate Domain Generation and Distribution Alignment for 2D Image-Based 3D Objects Retrieval	2127
. <i>Y. Su, Y. Li, D. Song, A. Liu, and J. Nie</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Blind Image Denoising via Dynamic Dual Learning	2139
. <i>Y. Du, G. Han, Y. Tan, C. Xiao, and S. He</i>	
<i>Multimodal Signal Representation and Visualization</i>	
Improving Generative Modelling in VAEs Using Multimodal Prior	2153
. <i>V. Abrol, P. Sharma, and A. Patra</i>	
<i>Video Surveillance and Semantic Analysis</i>	
STGL: Spatial-Temporal Graph Representation and Learning for Visual Tracking	2162
. <i>B. Jiang, Y. Zhang, B. Luo, X. Cao, and J. Tang</i>	
<i>Deep Learning for Multimedia Processing</i>	
Large Factor Image Super-Resolution With Cascaded Convolutional Neural Networks	2172
. <i>D. Zhang, J. Shao, Z. Liang, L. Gao, and H. T. Shen</i>	
<hr/>	
SPECIAL SECTION ON HYBRID HUMAN-ARTIFICIAL INTELLIGENCE FOR MULTIMEDIA COMPUTING	
Guest Editorial Special Section on Hybrid Human-Artificial Intelligence for Multimedia Computing	2185
. <i>R. Hamzaoui, H. Ning, C. Wang, R. Malekian, and W. Ding</i>	
<i>Video Surveillance and Semantic Analysis</i>	
Human Memory Update Strategy: A Multi-Layer Template Update Mechanism for Remote Visual Monitoring	2188
. <i>S. Liu, S. Wang, X. Liu, A. H. Gandomi, M. Daneshmand, K. Muhammad, and V. H. C. De Albuquerque</i>	
<i>Multimedia Search and Retrieval</i>	
Unsupervised Adversarial Instance-Level Image Retrieval	2199
. <i>C. Bai, H. Li, J. Zhang, L. Huang, and L. Zhang</i>	
<i>System Design and Optimization</i>	
Edge-Cloud Collaboration Enabled Video Service Enhancement: A Hybrid Human-Artificial Intelligence Scheme	2208
. <i>D. Wu, R. Bao, Z. Li, H. Wang, H. Zhang, and R. Wang</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Virtual Try-on Network With Attribute Transformation and Local Rendering	2222
. <i>J. Xu, Y. Pu, R. Nie, D. Xu, Z. Zhao, and W. Qian</i>	
<i>Compression and Coding</i>	
Virtual Try-on Network With Attribute Transformation and Local Rendering	2222
. <i>J. Xu, Y. Pu, R. Nie, D. Xu, Z. Zhao, and W. Qian</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Neural Style Palette: A Multimodal and Interactive Style Transfer From a Single Style Image	2245
. <i>J. J. Virtusio, J. J. M. Ople, D. S. Tan, M. Tanveer, N. Kumar, and K.-L. Hua</i>	

<i>Multimedia Search and Retrieval</i>	
Learning Low-Rank Sparse Representations With Robust Relationship Inference for Image Memorability Prediction	2259
. <i>P. Jing, Y. Shang, L. Nie, Y. Su, J. Liu, and M. Wang</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Enabling Artistic Control Over Pattern Density and Stroke Strength	2273
. <i>J. J. Virtusio, D. S. Tan, W.-H. Cheng, M. Tanveer, and K.-L. Hua</i>	
<i>Watermarking, Encryption, and Data Hiding</i>	
Separable Reversible Data Hiding for Encrypted Three-Dimensional Models Based on Spatial Subdivision and Space Encoding	2286
. <i>Y.-Y. Tsai</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Image Denoising Using Superpixel-Based PCA	2297
. <i>S. R. S. P. Malladi, S. Ram, and J. J. Rodríguez</i>	
Solving Jigsaw Puzzles via Nonconvex Quadratic Programming With the Projected Power Method	2310
. <i>F. Yan, Y. Zheng, J. Cong, L. Liu, D. Tao, and S. Hou</i>	
<i>Multimedia Search and Retrieval</i>	
Heterogeneous Community Question Answering via Social-Aware Multi-Modal Co-Attention Convolutional Matching	2321
. <i>J. Hu, S. Qian, Q. Fang, and C. Xu</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
HAPGN: Hierarchical Attentive Pooling Graph Network for Point Cloud Segmentation	2335
. <i>C. Chen, S. Qian, Q. Fang, and C. Xu</i>	
<i>Multimedia Search and Retrieval</i>	
Instance-Level Heterogeneous Domain Adaptation for Limited-Labeled Sketch-to-Photo Retrieval	2347
. <i>F. Yang, Y. Wu, Z. Wang, X. Li, S. Sakti, and S. Nakamura</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Toward Multi-Modal Conditioned Fashion Image Translation	2361
. <i>X. Gu, J. Yu, Y. Wong, and M. S. Kankanhalli</i>	
<i>Watermarking, Encryption, and Data Hiding</i>	
Universal Chosen-Ciphertext Attack for a Family of Image Encryption Schemes	2372
. <i>J. Chen, L. Chen, and Y. Zhou</i>	
<i>Multimedia Search and Retrieval</i>	
Learning Coarse-to-Fine Graph Neural Networks for Video-Text Retrieval	2386
. <i>W. Wang, J. Gao, X. Yang, and C. Xu</i>	
<i>Wireless and Mobile Multimedia</i>	
QoE-Driven UAV-Enabled Pseudo-Analog Wireless Video Broadcast: A Joint Optimization of Power and Trajectory	2398
. <i>X.-W. Tang, X.-L. Huang, and F. Hu</i>	
<i>Deep Learning for Multimedia Analysis</i>	
Fine-Grained Image Captioning With Global-Local Discriminative Objective	2413
. <i>J. Wu, T. Chen, H. Wu, Z. Yang, G. Luo, and L. Lin</i>	
<i>Deep Learning for Multimedia Processing</i>	
Joint Cross-Modal and Unimodal Features for RGB-D Salient Object Detection	2428
. <i>N. Huang, Y. Liu, Q. Zhang, and J. Han</i>	
<i>Multimedia Content Delivery Networks</i>	
Efficient Design and Control for Network-Assisted Device-to-Device Content Delivery Network	2442
. <i>J. Sung and D. Lee</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
GAC-GAN: A General Method for Appearance-Controllable Human Video Motion Transfer	2457
. <i>D. Wei, X. Xu, H. Shen, and K. Huang</i>	
<i>Deep Learning for Multimedia Analysis</i>	
SRD: A Tree Structure Based Decoder for Online Handwritten Mathematical Expression Recognition	2471
. <i>J. Zhang, J. Du, Y. Yang, Y.-Z. Song, and L. Dai</i>	
<i>Deep Learning for Multimedia Processing</i>	
Deep Single Image Deraining via Modeling Haze-Like Effect	2481
. <i>Y. Wang, D. Gong, J. Yang, Q. Shi, A. vanden Hengel, D. Xie, and B. Zeng</i>	

<i>Sparse Multimedia Signal Processing</i>	
Adaptive Graph Completion Based Incomplete Multi-View Clustering	2493
..... J. Wen, K. Yan, Z. Zhang, Y. Xu, J. Wang, L. Fei, and B. Zhang	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Estimation of Quality Scores From Subjective Tests-Beyond Subjects' MOS	2505
..... S. Pezzulli, M. G. Martini, and N. Barman	
<i>Deep Learning for Multimedia Analysis</i>	
Object-Aware Multimodal Named Entity Recognition in Social Media Posts With Adversarial Learning	2520
..... C. Zheng, Z. Wu, T. Wang, Y. Cai, and Q. Li	
<i>Media Conversion and Transcoding</i>	
YuvConv: Multi-Scale Non-Uniform Convolution Structure Based on YUV Color Model	2533
..... Y. Xiao, Z. Cai, and X. Yuan	
<i>Deep Learning for Multimedia Processing</i>	
Emotion Attention-Aware Collaborative Deep Reinforcement Learning for Image Cropping	2545
..... X. Zhang, Z. Li, and J. Jiang	
<i>Display Technology for Multimedia</i>	
Learning to Generate Multi-Exposure Stacks With Cycle Consistency for High Dynamic Range Imaging	2561
..... S. Lee, S. Y. Jo, G. H. An, and S.-J. Kang	
<i>Multimodal Human-machine Interfaces and Interaction</i>	
Robust CAPTCHAs Towards Malicious OCR	2575
..... J. Zhang, J. Sang, K. Xu, S. Wu, X. Zhao, Y. Sun, Y. Hu, and J. Yu	
<i>Mobile Media Technology and Systems</i>	
Data-Driven Bandwidth Prediction Models and Automated Model Selection for Low Latency	2588
..... A. Bentaleb, A. C. Begen, S. Harous, and R. Zimmermann	
<i>Watermarking, Encryption, and Data Hiding</i>	
Blind Image Clustering for Camera Source Identification via Row-Sparsity Optimization	2602
..... X. Jiang, S. Wei, T. Liu, R. Zhao, Y. Zhao, and H. Huang	
<i>Video Surveillance and Semantic Analysis</i>	
Group Re-Identification With Group Context Graph Neural Networks	2614
..... J. Zhu, H. Yang, W. Lin, N. Liu, J. Wang, and W. Zhang	
<i>Sparse Multimedia Signal Processing</i>	
Multi-Channel Deep Networks for Block-Based Image Compressive Sensing	2627
..... S. Zhou, Y. He, Y. Liu, C. Li, and J. Zhang	
<i>Audio/Speech/Language Analysis and Synthesis</i>	
Vector-Based Feature Representations for Speech Signals: From Supervector to Latent Vector	2641
..... Y. Jiang and F. H. F. Leung	
<i>Watermarking, Encryption, and Data Hiding</i>	
Robust Coding of Encrypted Images via 2D Compressed Sensing	2656
..... B. Zhang, D. Xiao, and Y. Xiang	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
AFNet: Temporal Locality-Aware Network With Dual Structure for Accurate and Fast Action Detection	2672
..... G. Chen, C. Zhang, and Y. Zou	
VehicleNet: Learning Robust Visual Representation for Vehicle Re-Identification	2683
..... Z. Zheng, T. Ruan, Y. Wei, Y. Yang, and T. Mei	
<i>Deep Learning for Multimedia Processing</i>	
Staged Sketch-to-Image Synthesis via Semi-supervised Generative Adversarial Networks	2694
..... Z. Li, C. Deng, E. Yang, and D. Tao	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Arbitrarily-Oriented Text Detection in Low Light Natural Scene Images	2706
..... M. Xue, P. Shivakumara, C. Zhang, Y. Xiao, T. Lu, U. Pal, D. Lopresti, and Z. Yang	

<i>Multimedia Search and Retrieval</i>	
Universal Cross-Domain 3D Model Retrieval	<i>D. Song, T.-B. Li, W.-H. Li, W.-Z. Nie, W. Liu, and A.-A. Liu</i> 2721
<i>Deep Learning for Multimedia Analysis</i>	
Adversarial Network With Multiple Classifiers for Open Set Domain Adaptation	<i>T. Shermin, G. Lu, S. W. Teng, M. Murshed, and F. Sohel</i> 2732
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Learning Specific and General Realm Feature Representations for Image Fusion	<i>F. Zhao and W. Zhao</i> 2745
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Quality Evaluation for Image Retargeting With Instance Semantics	<i>L. Li, Y. Li, J. Wu, L. Ma, and Y. Fang</i> 2757
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Deep Unsupervised Binary Descriptor Learning Through Locality Consistency and Self Distinctiveness	<i>B. Fan, H. Liu, H. Zeng, J. Zhang, X. Liu, and J. Han</i> 2770
<i>Multimodal Perception, Integration, and Multisensory Fusion</i>	
A Matrix Factorization Based Framework for Fusion of Physical and Social Sensors	<i>Y. Wang, F. Gelli, C. vonder Weth, and M. Kankanhalli</i> 2782
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Exploring the Representativity of Art Paintings	<i>Y. Deng, F. Tang, W. Dong, C. Ma, F. Huang, O. Deussen, and C. Xu</i> 2794
<i>Compression and Coding</i>	
Efficient Projected Frame Padding for Video-Based Point Cloud Compression	<i>L. Li, Z. Li, S. Liu, and H. Li</i> 2806
<i>Virtual and Augmented Reality</i>	
An Accurate, Robust Visual Odometry and Detail-Preserving Reconstruction System	<i>X. Gong, Y. Liu, Q. Wu, J. Huang, H. Zong, and J. Wang</i> 2820
<i>Multimedia Search and Retrieval</i>	
StyleGuide: Zero-Shot Sketch-Based Image Retrieval Using Style-Guided Image Generation	<i>T. Dutta, A. Singh, and S. Biswas</i> 2833
<i>Image/Video/Graphics Analysis and Synthesis</i>	
DLGAN: Depth-Preserving Latent Generative Adversarial Network for 3D Reconstruction	<i>C. Liu, D. Kong, S. Wang, J. Li, and B. Yin</i> 2843
<i>Compression and Coding</i>	
Hardness-Aware Dictionary Learning: Boosting Dictionary for Recognition	<i>L. Wang, S. Li, S. Wang, D. Kong, and B. Yin</i> 2857
Fine Granularity Access in Interactive Compression of 360-Degree Images Based on Rate-adaptive Channel Codes	<i>N. Mahmoudian Bidgoli, T. Maugey, and A. Roumy</i> 2868
<i>Image/Video/Graphics Analysis and Synthesis</i>	
An Adaptive Arbitrary Multiresolution Decomposition for Multiscale Geometric Analysis	<i>Z. Lu, G. Yang, J. Yang, and Y. Wang</i> 2883
<i>Multimedia for Social Activities in Health, Environment, and Urban Living</i>	
Learning Diverse Fashion Collocation by Neural Graph Filtering	<i>X. Liu, Y. Sun, Z. Liu, and D. Lin</i> 2894
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Saliency Detection Using Deep Features and Affinity-Based Robust Background Subtraction	<i>M. Nawaz and H. Yan</i> 2902
Hybrid-Attention Enhanced Two-Stream Fusion Network for Video Venue Prediction	<i>Y. Zhang, W. Min, L. Nie, and S. Jiang</i> 2917
Learning Compact Multifeature Codes for Palmprint Recognition From a Single Training Image per Palm	<i>L. Fei, B. Zhang, L. Zhang, W. Jia, J. Wen, and J. Wu</i> 2930
<i>Sparse Multimedia Signal Processing</i>	
Unsupervised Multi-View Clustering by Squeezing Hybrid Knowledge From Cross View and Each View	<i>J. Tan, Y. Shi, Z. Yang, C. Wen, and L. Lin</i> 2943

ICME 2020 BEST PAPERS

<i>Compression and Coding</i>	
Towards Coding for Human and Machine Vision: Scalable Face Image Coding	2957
. S. Yang, Y. Hu, W. Yang, L.-Y. Duan, and J. Liu	
Light Field Image Coding Using VVC Standard and View Synthesis Based on Dual Discriminator GAN	2972
. N. Bakir, W. Hamidouche, S. A. Fezza, K. Samrouth, and O. Déforges	
<i>Deep Learning for Multimedia Processing</i>	
Motion Blur Removal With Quality Assessment Guidance	2986
. J. Li, B. Yan, Q. Lin, A. Li, and C. Ma	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Expression-Aware Face Reconstruction via a Dual-Stream Network	2998
. X. Chai, J. Chen, C. Liang, D. Xu, and C.-W. Lin	
<i>Compression and Coding</i>	
Learned Multi-Resolution Variable-Rate Image Compression With Octave-Based Residual Blocks	3013
. M. Akbari, J. Liang, J. Han, and C. Tu	
<i>3-D Video Signal Processing</i>	
Dynamic Point Cloud Inpainting via Spatial-Temporal Graph Learning	3022
. Z. Fu and W. Hu	
<i>Audio/Speech/Language Analysis and Synthesis</i>	
Intelligibility Enhancement Via Normal-to-Lombard Speech Conversion With Long Short-Term Memory Network and Bayesian Gaussian Mixture Model	3035
. G. Li, X. Wang, R. Hu, H. Zhang, and S. Ke	
<i>Deep Learning for Multimedia Analysis</i>	
Spatial Pyramid Attention for Deep Convolutional Neural Networks	3048
. X. Ma, J. Guo, A. Sansom, M. McGuire, A. Kalaani, Q. Chen, S. Tang, Q. Yang, and S. Fu	
<hr/>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Attentive Composite Residual Network for Robust Rain Removal from Single Images	3059
. Y. Que, S. Li, and H. J. Lee	
<i>Multimedia Storytelling and Cross-Modal Translations Between Multimedia Contents</i>	
Interclass-Relativity-Adaptive Metric Learning for Cross-Modal Matching and Beyond	3073
. F. Chen, J. Shao, Y. Zhang, X. Xu, and H. T. Shen	
<i>Deep Learning for Multimedia Analysis</i>	
Attribute-Aware Pedestrian Detection in a Crowd	3085
. J. Zhang, L. Lin, J. Zhu, Y. Li, Y.-chen Chen, Y. Hu, and S. C. H. Hoi	
<i>3-D Processing and Presentation</i>	
Mesh Convolution: A Novel Feature Extraction Method for 3D Nonrigid Object Classification	3098
. Y. Chen, J. Zhao, C. Shi, and D. Yuan	
<i>Deep Learning for Multimedia Analysis</i>	
Learning to Segment Video Object With Accurate Boundaries	3112
. J. Cheng, Y. Yuan, Y. Li, J. Wang, and S. Wang	
Graph Regularized Encoder-Decoder Networks for Image Representation Learning	3124
. S. Yang, L. Li, S. Wang, W. Zhang, Q. Huang, and Q. Tian	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Latent Representation Learning Model for Multi-Band Images Fusion via Low-Rank and Sparse Embedding	3137
. B. Wang, H. Niu, J. Zeng, G. Bai, S. Lin, and Y. Wang	
<i>Deep Learning for Multimedia Processing</i>	
Luminance-Aware Pyramid Network for Low-Light Image Enhancement	3153
. J. Li, J. Li, F. Fang, F. Li, and G. Zhang	
<i>System Design and Optimization</i>	
Optimizing Video QoE for Mobile eMBMS Users in Cellular Networks	3166
. A. Khalid, A. H. Zahran, and C. J. Sreenan	

<i>Watermarking, Encryption, and Data Hiding</i>	
Frame-Wise Detection of Double HEVC Compression by Learning Deep Spatio-Temporal Representations in Compression Domain	P. He, H. Li, H. Wang, S. Wang, X. Jiang, and R. Zhang 3179
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Co-Saliency Detection via a General Optimization Model and Adaptive Graph Learning	B. Jiang, X. Jiang, J. Tang, and B. Luo 3193
<i>Knowledge and Semantics Modeling for Multimedia Databases</i>	
Unsupervised Video Summarization via Relation-Aware Assignment Learning	J. Gao, X. Yang, Y. Zhang, and C. Xu 3203
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Handling Outliers by Robust M-Estimation in Blind Image Deblurring	X. Zhang, R. Wang, D. Chen, Y. Zhao, and W. Gao 3215
<i>Multimedia for Augmented Experience in Real and Virtual Life</i>	
Attention-Based Deep Reinforcement Learning for Virtual Cinematography of 360° Videos	J. Wang, M. Xu, L. Jiang, and Y. Song 3227
<i>Multimodal Perception, Integration, and Multisensory Fusion</i>	
Global Manifold Learning for Interactive Image Segmentation	T. Wang, Z. Ji, J. Yang, Q. Sun, and P. Fu 3239
<i>Green Multimedia Communications and Computing</i>	
NDN-MMRA: Multi-Stage Multicast Rate Adaptation in Named Data Networking WLAN	F. Wu, W. Yang, J. Ren, F. Lyu, P. Yang, Y. Zhang, and X. Shen 3250
<i>Deep Learning for Multimedia Analysis</i>	
Adaptive and Robust Partition Learning for Person Retrieval With Policy Gradient	Y. Shi, Z. Wei, H. Ling, Z. Wang, P. Zhu, J. Shen, and P. Li 3264
<i>Compression and Coding</i>	
Model-Based Joint Bit Allocation Between Geometry and Color for Video-Based 3D Point Cloud Compression	Q. Liu, H. Yuan, J. Hou, R. Hamzaoui, and H. Su 3278
<i>Video Surveillance and Semantic Analysis</i>	
Adversarial 3D Convolutional Auto-Encoder for Abnormal Event Detection in Videos	C. Sun, Y. Jia, H. Song, and Y. Wu 3292
<i>Deep Learning for Multimedia Analysis</i>	
Temporal Textual Localization in Video via Adversarial Bi-Directional Interaction Networks	Z. Zhang, Z. Zhao, Z. Zhang, Z. Lin, Q. Wang, and R. Hong 3306
<i>Image/Video/Graphics Analysis and Synthesis</i>	
KTransGAN: Variational Inference-Based Knowledge Transfer for Unsupervised Conditional Generative Learning	M. Azzam, W. Wu, W. Cao, S. Wu, and H.-S. Wong 3318
<i>Video Surveillance and Semantic Analysis</i>	
DCR: A Unified Framework for Holistic/Partial Person ReID	Z. Gao, L. Gao, H. Zhang, Z. Cheng, R. Hong, and S. Chen 3332
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Temporal Constraint Background-Aware Correlation Filter With Saliency Map	J. Liao, C. Qi, and J. Cao 3346
<i>Deep Learning for Multimedia Processing</i>	
PFAN++: Bi-Directional Image-Text Retrieval With Position Focused Attention Network	Y. Wang, H. Yang, X. Bai, X. Qian, L. Ma, J. Lu, B. Li, and X. Fan 3362
<i>Audio/Speech/Language Analysis and Synthesis</i>	
Speaker Clustering by Co-Optimizing Deep Representation Learning and Cluster Estimation	Y. Li, W. Wang, M. Liu, Z. Jiang, and Q. He 3377
<i>3-D Processing and Presentation</i>	
Salient Object Detection in Stereoscopic 3D Images Using a Deep Convolutional Residual Autoencoder	W. Zhou, J. Wu, J. Lei, J.-N. Hwang, and L. Yu 3388

<i>Deep Learning for Multimedia Analysis</i>	
Deep Unsupervised Self-Evolutionary Hashing for Image Retrieval	3400
. <i>H. Zhang, Y. Gu, Y. Yao, Z. Zhang, L. Liu, J. Zhang, and L. Shao</i>	
<i>Audio/Speech/Language Analysis and Synthesis</i>	
Speech Personality Recognition Based on Annotation Classification Using Log-Likelihood Distance and Extraction of Essential Audio Features	3414
. <i>Z.-T. Liu, A. Rehman, M. Wu, W.-H. Cao, and M. Hao</i>	
<i>Deep Learning for Multimedia Processing</i>	
LCSegNet: An Efficient Semantic Segmentation Network for Large-Scale Complex Chinese Character Recognition	3427
. <i>X. Wu, Q. Chen, Y. Xiao, W. Li, X. Liu, and B. Hu</i>	
Universal-to-Specific Framework for Complex Action Recognition	3441
. <i>P. Zhao, L. Xie, Y. Zhang, and Q. Tian</i>	
Explore Video Clip Order With Self-Supervised and Curriculum Learning for Video Applications	3454
. <i>J. Xiao, L. Li, D. Xu, C. Long, J. Shao, S. Zhang, S. Pu, and Y. Zhuang</i>	
<i>Watermarking, Encryption, and Data Hiding</i>	
Blind 3D-Printing Watermarking Using Moment Alignment and Surface Norm Distribution	3467
. <i>A. Delmotte, K. Tanaka, H. Kubo, T. Funatomi, and Y. Mukaigawa</i>	
<i>Deep Learning for Multimedia Analysis</i>	
Deep Multi-View Subspace Clustering With Unified and Discriminative Learning	3483
. <i>Q. Wang, J. Cheng, Q. Gao, G. Zhao, and L. Jiao</i>	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Improving Student Learning Satisfaction by Using an Innovative DASH-Based Multiple Sensorial Media Delivery Solution	3494
. <i>T. Bi, R. Lyons, G. Fox, and G.-M. Muntean</i>	
<i>Watermarking, Encryption, and Data Hiding</i>	
A Serial Image Copy-Move Forgery Localization Scheme With Source/Target Distinguishment	3506
. <i>B. Chen, W. Tan, G. Coatrieux, Y. Zheng, and Y.-Q. Shi</i>	
<i>Deep Learning for Multimedia Analysis</i>	
Visual Question Answering With Dense Inter- and Intra-Modality Interactions	3518
. <i>F. Liu, J. Liu, Z. Fang, R. Hong, and H. Lu</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Self-Supervised Deep TripleNet for Video Object Segmentation	3530
. <i>K. Xu, L. Wen, G. Li, and Q. Huang</i>	
<i>Multimodal Human-Machine Interfaces and Interaction</i>	
Monocular 3D Facial Expression Features for Continuous Affect Recognition	3540
. <i>E. Pei, M. C. Oveneke, Y. Zhao, D. Jiang, and H. Sahli</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
MVANet: Multi-Task Guided Multi-View Attention Network for Chinese Food Recognition	3551
. <i>H. Liang, G. Wen, Y. Hu, M. Luo, P. Yang, and Y. Xu</i>	
<i>Video Surveillance and Semantic Analysis</i>	
Learning Spatial-Temporal Representations Over Walking Tracklet for Long-Term Person Re-Identification in the Wild	3562
. <i>P. Zhang, J. Xu, Q. Wu, Y. Huang, and X. Ben</i>	
<i>Low-Complexity Processing in Devices and Sensors</i>	
A New Multihypothesis-Based Compressed Video Sensing Reconstruction System	3577
. <i>S. Zheng, J. Chen, X.-P. Zhang, and Y. Kuo</i>	
<i>Multimedia Search and Retrieval</i>	
Deep Semantic Parsing of Freehand Sketches With Homogeneous Transformation, Soft-Weighted Loss, and Staged Learning	3590
. <i>Y. Zheng, H. Yao, and X. Sun</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Wildfish++: A Comprehensive Fish Benchmark for Multimedia Research	3603
. <i>P. Zhuang, Y. Wang, and Y. Qiao</i>	

<i>Subjective and Objective Quality Assessment and User Experience</i>	
Learning the Relation Between Interested Objects and Aesthetic Region for Image Cropping	3618
. <i>P. Lu, H. Zhang, X. Peng, and X. Jin</i>	
<i>Multimedia Streaming and Transport</i>	
CoLEAP: Cooperative Learning-Based Edge Scheme With Caching and Prefetching for DASH Video Delivery	3631
. <i>W. Shi, C. Wang, Y. Jiang, Q. Li, G. Shen, and G.-M. Muntean</i>	
<i>Multimedia Search and Retrieval</i>	
Collaborative Image Relevance Learning for Visual Re-Ranking	3646
. <i>J. Ouyang, W. Zhou, M. Wang, Q. Tian, and H. Li</i>	
<i>Deep Learning for Multimedia Processing</i>	
Parametric Shape Estimation of Human Body Under Wide Clothing	3657
. <i>Y. Lu, J.-H. Cha, S.-K. Youm, and S.-W. Jung</i>	
<i>Wireless and Mobile Multimedia</i>	
Optimal Wireless Streaming of Multi-Quality 360 VR Video By Exploiting Natural, Relative Smoothness-Enabled, and Transcoding-Enabled Multicast Opportunities	3670
. <i>K. Long, Y. Cui, C. Ye, and Z. Liu</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Privacy-Preserving In-Home Fall Detection Using Visual Shielding Sensing and Private Information-Embedding	3684
. <i>J. Liu, R. Tan, G. Han, N. Sun, and S. Kwong</i>	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Comparative Perceptual Assessment of Visual Signals Using Free Energy Features	3700
. <i>G. Zhai, Y. Zhu, and X. Min</i>	
<i>Multimedia Content Delivery Networks</i>	
An Approximation Algorithm to Maximize User Capacity for an Auto-Scaling VoD System	3714
. <i>Z. Chang and S.-H. G. Chan</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
One-Shot Texture Retrieval Using Global Grouping Metric	3726
. <i>K. Zhu, Y. Cao, W. Zhai, and Z.-J. Zha</i>	
<i>Multimodal Perception, Integration, and Multisensory Fusion</i>	
Hard Pixel Mining for Depth Privileged Semantic Segmentation	3738
. <i>Z. Gu, L. Niu, H. Zhao, and L. Zhang</i>	
<i>Military, Security, and Intelligence</i>	
Low-Cost Anti-Copying 2D Barcode by Exploiting Channel Noise Characteristics	3752
. <i>N. Xie, Q. Zhang, Y. Chen, J. Hu, G. Luo, and C. Chen</i>	
<i>Media Conversion and Transcoding</i>	
A New Approach for Character Recognition of Multi-Style Vehicle License Plates	3768
. <i>Q. Huang, Z. Cai, and T. Lan</i>	
<i>Multimedia Search and Retrieval</i>	
Deep Hashing With Weighted Spatial Importance	3778
. <i>Y. Shi, X. Nie, M. Chen, L. Lian, and Y. Yin</i>	
<i>3-D Video Signal Processing</i>	
C-GCN: Correlation Based Graph Convolutional Network for Audio-Video Emotion Recognition	3793
. <i>W. Nie, M. Ren, J. Nie, and S. Zhao</i>	
<i>Multimedia Content Delivery Networks</i>	
Controlling P2P-CDN Live Streaming Services at SDN-Enabled Multi-Access Edge Datacenters	3805
. <i>S. Nacakli and A. M. Tekalp</i>	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Wide Color Gamut Image Content Characterization: Method, Evaluation, and Applications	3817
. <i>J. Lee, T. Vigier, P. Le Callet, and J.-S. Lee</i>	
<i>Multimodal Perception, Integration, and Multisensory Fusion</i>	
Kernelized Multiview Subspace Analysis By Self-Weighted Learning	3828
. <i>H. Wang, Y. Wang, Z. Zhang, X. Fu, L. Zhuo, M. Xu, and M. Wang</i>	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Adaptive Multi-Feature Reliability Re-Determinative Correlation Filter for Visual Tracking	3841
. <i>M. Guan and C. Wen</i>	

<i>Immersive Communications and Networking</i>	
Sparkle: User-Aware Viewport Prediction in 360-Degree Video Streaming	3853
..... J. Chen, X. Luo, M. Hu, D. Wu, and Y. Zhou	
<i>Compression and Coding</i>	
String Prediction for 4:2:0 Format Screen Content Coding and Its Implementation in AVS3	3867
..... Q. Zhou, L. Zhao, K. Zhou, T. Lin, H. Wang, S. Wang, and M. Jiao	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Predicting the Perceptual Quality of Point Cloud: A 3D-to-2D Projection-Based Exploration	3877
..... Q. Yang, H. Chen, Z. Ma, Y. Xu, R. Tang, and J. Sun	
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>	
Hierarchical Group-Level Emotion Recognition	3892
..... K. Fujii, D. Sugimura, and T. Hamamoto	
<i>Multimedia Search and Retrieval</i>	
Learning to Hash With Dimension Analysis Based Quantizer for Image Retrieval	3907
..... Y. Cao, H. Qi, J. Gui, K. Li, Y. Y. Tang, and J. T.-Y. Kwok	
<i>Deep Learning for Multimedia Analysis</i>	
Domain-Oriented Semantic Embedding for Zero-Shot Learning	3919
..... S. Min, H. Yao, H. Xie, Z.-J. Zha, and Y. Zhang	
<i>Big Data Analytics on Multimedia Data and Crowd Sourcing for Multimedia Applications</i>	
Anisotropic Graph Convolutional Network for Semi-Supervised Learning	3931
..... M. Mesgaran and A. B. Hamza	
<i>Multimedia Search and Retrieval</i>	
Rank-Consistency Deep Hashing for Scalable Multi-Label Image Search	3943
..... C. Ma, J. Lu, and J. Zhou	
User Identity Linkage Across Social Media via Attentive Time-Aware User Modeling	3957
..... X. Chen, X. Song, S. Cui, T. Gan, Z. Cheng, and L. Nie	
Aggregating Global and Local Visual Representation for Vehicle Re-Identification	3968
..... X. Lin, R. Li, X. Zheng, P. Peng, Y. Wu, F. Huang, and R. Ji	
<i>Audio/Speech/Language Analysis and Synthesis</i>	
Environmental Sound Classification Using Local Binary Pattern and Audio Features Collaboration	3978
..... O. K. Toffa and M. Mignotte	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Dynamic Motion Estimation and Evolution Video Prediction Network	3986
..... N. Kim and J.-W. Kang	
<i>Knowledge and Semantics Modeling for Multimedia Databases</i>	
Emotion Knowledge Driven Video Highlight Detection	3999
..... F. Qi, X. Yang, and C. Xu	
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>	
Image-Text Multimodal Emotion Classification via Multi-View Attentional Network	4014
..... X. Yang, S. Feng, D. Wang, and Y. Zhang	
<i>Multimedia Storytelling and Cross-Modal Translations Between Multimedia Contents</i>	
Story-driven Video Editing	4027
..... Z. Wang, J. Li, and Y.-G. Jiang	
<i>Subjective and Objective Quality Assessment and User Experience</i>	
MaD-DLS: Mean and Deviation of Deep and Local Similarity for Image Quality Assessment	4037
..... K. Sim, J. Yang, W. Lu, and X. Gao	
Point Cloud Rendering After Coding: Impacts on Subjective and Objective Quality	4049
..... A. Javaheri, C. Brites, F. Pereira, and J. Ascenso	
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Pixel-Level Non-local Image Smoothing With Objective Evaluation	4065
..... J. Xu, Z.-A. Liu, Y.-K. Hou, X.-T. Zhen, L. Shao, and M.-M. Cheng	
<i>Multimedia Search and Retrieval</i>	
Adaptive Partial Multi-View Hashing for Efficient Social Image Retrieval	4079
..... C. Zheng, L. Zhu, Z. Cheng, J. Li, and A.-A. Liu	

TBEFN: A Two-Branch Exposure-Fusion Network for Low-Light Image Enhancement	<i>K. Lu and L. Zhang</i>	4093
<i>Video Surveillance and Semantic Analysis</i>		
Multi-Encoder Towards Effective Anomaly Detection in Videos	<i>Z. Fang, J. T. Zhou, Y. Xiao, Y. Li, and F. Yang</i>	4106
<i>3-D Processing and Presentation</i>		
Acoustic Room Modelling Using 360 Stereo Cameras	<i>H. Kim, L. Remaggi, S. Fowler, P. JB Jackson, and A. Hilton</i>	4117
<i>Audio/Speech/Language Analysis and Synthesis</i>		
CAA-Net: Conditional Atrous CNNs With Attention for Explainable Device-Robust Acoustic Scene Classification	<i>Z. Ren, Q. Kong, J. Han, M. D. Plumbley, and B. W. Schuller</i>	4131
<i>Image/Video/Graphics Analysis and Synthesis</i>		
Modeling Fashion Influence From Photos	<i>Z. Al-Halah and K. Grauman</i>	4143
Iterative Knowledge Distillation for Automatic Check-Out	<i>L. Zhang, D. Du, C. Li, Y. Wu, and T. Luo</i>	4158
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>		
Transformer Encoder With Multi-Modal Multi-Head Attention for Continuous Affect Recognition	<i>H. Chen, D. Jiang, and H. Sahli</i>	4171
<i>Image/Video/Graphics Analysis and Synthesis</i>		
A Resource-Efficient Parallel Connected Component Labeling Algorithm and Its Hardware Implementation	<i>C. Zhao, W. Gao, and F. Nie</i>	4184
<i>Transport and Automotive</i>		
Improving Driver Gaze Prediction With Reinforced Attention	<i>K. Lv, H. Sheng, Z. Xiong, W. Li, and L. Zheng</i>	4198
<i>System Design and Optimization</i>		
Robust and Efficient RGB-D SLAM in Dynamic Environments	<i>X. Yang, Z. Yuan, D. Zhu, C. Chi, K. Li, and C. Liao</i>	4208
<i>Healthcare</i>		
Domain Adaptation for Food Intake Classification With Teacher/Student Learning	<i>M. A. T. Turan and E. Erzin</i>	4220
<i>Image/Video/Graphics Analysis and Synthesis</i>		
Capturing Relevant Context for Visual Tracking	<i>Y. Zhang, B. Ma, J. Wu, L. Huang, and J. Shen</i>	4232
<i>Subjective and Objective Quality Assessment and User Experience</i>		
Re-Visiting Discriminator for Blind Free-Viewpoint Image Quality Assessment	<i>S. Ling, J. Li, Z. Che, W. Zhou, J. Wang, and P. Le Callet</i>	4245
Blind Quality Assessment of Screen Content Images Via Macro-Micro Modeling of Tensor Domain Dictionary	<i>Y. Bai, Z. Zhu, G. Jiang, and H. Sun</i>	4259
<i>Image/Video/Graphics Analysis and Synthesis</i>		
DSLRL: Deep Stacked Laplacian Restorer for Low-Light Image Enhancement	<i>S. Lim and W. Kim</i>	4272
<i>Multimedia Search and Retrieval</i>		
Learning Dual-Pooling Graph Neural Networks for Few-Shot Video Classification	<i>Y. Hu, J. Gao, and C. Xu</i>	4285
<i>Deep Learning for Multimedia Analysis</i>		
Discriminative Region Mining for Object Detection	<i>L. Chen, H. Zheng, Z. Yan, and Y. Li</i>	4297
<i>Media Cloud Computing and Communications</i>		
Collaborative Social-Aware and QoE-Driven Video Caching and Adaptation in Edge Network	<i>Y. Chiang, C.-H. Hsu, and H.-Y. Wei</i>	4311
<i>Subjective and Objective Quality Assessment and User Experience</i>		
TTL-IQA: Transitive Transfer Learning Based No-Reference Image Quality Assessment	<i>X. Yang, F. Li, and H. Liu</i>	4326
<i>Compression and Coding</i>		
Redundancy and Optimization of tANS Entropy Encoders	<i>I. Blanes, M. Hernández-Cabronero, J. Serra-Sagrìstà, and M. W. Marcellin</i>	4341
<i>Multimedia Search and Retrieval</i>		
SEA: Sentence Encoder Assembly for Video Retrieval by Textual Queries	<i>X. Li, F. Zhou, C. Xu, J. Ji, and G. Yang</i>	4351

<i>Image/Video/Graphics Analysis and Synthesis</i>	
Temporal Action Localization Using Long Short-Term Dependency	Y. Zhou, R. Wang, H. Li, and S.-Y. Kung 4363
<i>Deep Learning for Multimedia Analysis</i>	
Person Retrieval in Surveillance Videos Via Deep Attribute Mining and Reasoning	Y. Shi, Z. Wei, H. Ling, Z. Wang, J. Shen, and P. Li 4376
<i>Compression and Coding</i>	
Fast Multi-Type Tree Partitioning for Versatile Video Coding Using a Lightweight Neural Network	S.-hyo Park and J.-W. Kang 4388
<i>Image/Video/Graphics Analysis and Synthesis</i>	
Learning Fundamental Visual Concepts Based on Evolved Multi-Edge Concept Graph	H. Wang, Y. Du, G. Zhang, Z. Cai, and C. Su 4400
Parameter Sharing Exploration and Hetero-Center Triplet Loss for Visible-Thermal Person Re-Identification	H. Liu, X. Tan, and X. Zhou 4414
<i>Deep Learning for Multimedia Analysis</i>	
Referring Expression Comprehension: A Survey of Methods and Datasets	Y. Qiao, C. Deng, and Q. Wu 4426
<i>Knowledge and Semantics Modeling for Multimedia Databases</i>	
Multimodal Disentangled Domain Adaption for Social Media Event Rumor Detection	H. Zhang, S. Qian, Q. Fang, and C. Xu 4441
<i>Compression and Coding</i>	
PVC-SLP: Perceptual Vibrotactile-Signal Compression Based-on Sparse Linear Prediction	R. Hassen, B. Gülecyüz, and E. Steinbach 4455
<i>Multimedia Sentiment Analysis and Synthesis; Affective Media Processing</i>	
APSE: Attention-Aware Polarity-Sensitive Embedding for Emotion-Based Image Retrieval	X. Yao, S. Zhao, Y.-K. Lai, D. She, J. Liang, and J. Yang 4469
<i>Image/Video/Graphics Analysis and Synthesis</i>	
From Edge to Keypoint: An End-to-End Framework For Indoor Layout Estimation	W. Zhang, Q. Zhang, W. Zhang, J. Gu, and Y. Li 4483
<i>Deep Learning for Multimedia Analysis</i>	
A Hybrid Approach for Detecting Prerequisite Relations in Multi-Modal Food Recipes	L. Pan, J. Chen, S. Liu, C.-W. Ngo, M.-Y. Kan, and T.-S. Chua 4491
<i>Deep Learning for Multimedia Processing</i>	
Deep Texture Exemplar Extraction Based on Trimmed T-CNN	H. Wu, W. Yan, P. Li, and Z. Wen 4502
<i>Knowledge and Semantics Modeling for Multimedia Databases</i>	
Adaptively Clustering-Driven Learning for Visual Relationship Detection	A.-A. Liu, Y. Wang, N. Xu, W. Nie, J. Nie, and Y. Zhang 4515
<i>Subjective and Objective Quality Assessment and User Experience</i>	
Predicting User Quitting Ratio in Adaptive Bitrate Video Streaming	P. Lebreton and K. Yamagishi 4526
<i>Multimedia Search and Retrieval</i>	
Semantic-Driven Interpretable Deep Multi-Modal Hashing for Large-Scale Multimedia Retrieval	X. Lu, L. Liu, L. Nie, X. Chang, and H. Zhang 4541
<i>Sparse Multimedia Signal Processing</i>	
On Reliable Multi-View Affinity Learning for Subspace Clustering	X. Xiao, Y.-J. Gong, Z. Hua, and W.-N. Chen 4555
List of Reviewers	4567
2021 INDEX	Available online at http://ieeexplore.ieee.org
