Teaser Session W1

Paper id	Paper title
mmfp2127	Joint Local Relational Augmentation and Global Nash Equilibrium for Federated Learning with Non-IID Data
mmfp2142	FlexIcon: Flexible Icon Colorization via Guided Images and Palettes
mmfp2148	ECENet: Explainable and Context-Enhanced Network for Muti-modal Fact verification
mmfp2292	Aesthetics-Driven Virtual Time-Lapse Photography Generation
mmfp2416	A Blind Streaming System for Multi-client Online 6-DoF View Touring
mmfp2515	Elucidate Gender Fairness in Singing Voice Transcription
mmfp2568	On the Performance of Subjective Visual Quality Assessment Protocols for Nearly Visually Lossless Image Compression
mmfp2605	Relative NN-Descent: A Fast Index Construction for Graph-Based Approximate Nearest Neighbor Search
mmfp2809	IGG: Improved Graph Generation for Domain Adaptive Object Detection
mmfp3036	Draw2Edit: Mask-Free Sketch-Guided Image Manipulation
mmfp3103	ScaleFlow: Efficient Deep Vision Pipeline with Closed-Loop Scale-Adaptive Inference
mmfp3112	ENTRO: Tackling the Encoding and Networking Trade-off in Offloaded Video Analytics
mmfp3236	FFNeRV: Flow-Guided Frame-Wise Neural Representations for Videos
mmfp3328	Real-time Facial Animation for 3D Stylized Character with Emotion Dynamics
mmfp3453	VQBA: Visual-Quality-Driven Bit Allocation for Low-Latency Point Cloud Streaming
mmfp3862	Double Doodles: Sketching Animation in Immersive Environment With 3+6 DOFs Motion Gestures
mmfp3944	Reconnecting the Broken Civilization: Patchwork Integration of Fragments from Ancient Manuscripts
mmfp4034	A Symbolic Characters Aware Model for Solving Geometry Problems
mmfp4071	SGDiff: A Style Guided Diffusion Model for Fashion Synthesis
mmfp4091	Modeling Multi-Relational Connectivity for Personalized Fashion Matching

mmfp2425	TTPOINT: A Tensorized Point Cloud Network for Lightweight Action Recognition with Event Cameras
mmfp2663	Learning Pixel-wise Alignment for Unsupervised Image Stitching
mmfp2734	Parameter Exchange for Robust Dynamic Domain Generalization
mmfp2756	YOGA: Yet Another Geometry-based Point Cloud Compressor
mmfp2757	G-PCC++: Enhanced Geometry-based Point Cloud Compression
mmfp2958	DFIL: Deepfake Incremental Learning by Exploiting Domain-invariant Forgery Clues
mmfp2995	FlatGAN: A Holistic Approach for Robust Flat-Coloring in High-Definition with Understanding Line Discontinuity
mmfp3354	A Four-Pronged Defense Against Byzantine Attacks in Federated Learning
mmfp3469	Control3D: Towards Controllable Text-to-3D Generation
mmfp4080	Feeling Present! From Physical to Virtual Cinematography Lighting Education with Metashadow
mmfp4144	LHNet: A Low-cost Hybrid Network for Single Image Dehazing
mmfp4151	Learning Profitable NFT Image Diffusions via Multiple Visual-Policy Guided Reinforcement Learning

mmfp4216	Simple Techniques are Sufficient for Boosting Adversarial Transferability
mmfp4261	Gradient-Free Textual Inversion

All papers are presented also in the Poster Session W

Teaser Session W2

Paper id	Paper title
mmfp0048	What2comm: Towards Communication-efficient Collaborative Perception via Feature Decoupling
mmfp0123	Automatic Generation of Commercial Scenes
mmfp0154	CATR: Combinatorial-Dependence Audio-Queried Transformer for Audio-Visual Video Segmentation
mmfp0224	Finding Efficient Pruned Network via Refined Gradients for Pruned Weights
mmfp0356	Optimizing Adaptive Video Streaming with Human Feedback
mmfp0459	RoomDreamer: Text-Driven 3D Indoor Scene Synthesis with Coherent Geometry and Texture
mmfp0500	Cal-SFDA: Source-Free Domain-adaptive Semantic Segmentation with Differentiable Expected Calibration Error
mmfp1117	Pagoda: Privacy Protection for Volumetric Video Streaming through Poisson Diffusion Model
mmfp1189	Neural Video Compression with Spatio-Temporal Cross-Covariance Transformers
mmfp1225	GrooveMeter: Enabling Music Engagement-aware Apps by Detecting Reactions to Daily Music Listening via Earable Sensing
mmfp1348	ColSLAM: A Versatile Collaborative SLAM System for Mobile Phones Using Point-Line Features and Map Caching
mmfp1367	Mask-Guided Progressive Network for Joint Raindrop and Rain Streak Removal in Videos
mmfp1501	Adversarial Attack for Robust Watermark Protection Against Inpainting-based and Blind Watermark Removers
mmfp1773	The Effects of Viewing Formats and Song Genres on Audience Experiences in Virtual Avatar Concerts
mmfp1830	MetaCast: A Self-Driven Metaverse Announcer Architecture Based on Quality of Experience Evaluation Model
mmfp1932	LaDI-VTON: Latent Diffusion Textual-Inversion Enhanced Virtual Try-On
mmfp1949	An Order-Complexity Aesthetic Assessment Model for Aesthetic-aware Music Recommendation
mmfp2029	NIF: A Fast Implicit Image Compression with Bottleneck Layers and Modulated Sinusoidal Activations

Paper id	Paper title
mmfp0051	Multispectral Object Detection via Cross-Modal Conflict-Aware Learning
mmfp0289	Client-Adaptive Cross-Model Reconstruction Network for Modality-Incomplete Multimodal Federated Learning
mmfp0314	LGViT: Dynamic Early Exiting for Accelerating Vision Transformer
mmfp0389	FedGH: Heterogeneous Federated Learning with Generalized Global Header
mmfp0710	Toward Scalable Image Feature Compression: A Content-Adaptive and Diffusion-Based Approach
mmfp0941	Brighten-and-Colorize: A Decoupled Network for Customized Low-Light Image Enhancement
mmfp0947	FourLLIE: Boosting Low-Light Image Enhancement by Fourier Frequency Information
mmfp1027	Latent-space Unfolding for MRI Reconstruction
mmfp1259	Enhanced Image Deblurring: An Efficient Frequency Exploitation and Preservation Network
mmfp1539	ICMH-Net: Neural Image Compression Towards both Machine Vision and Human Vision
mmfp1571	A Reference-free Self-supervised Domain Adaptation Framework for Low-quality Fundus Image Enhancement
mmfp1722	Physical Adversarial Attack on Near-Infrared based Human Detector for Nighttime Surveillance Camera Systems
mmfp1755	Self-supervised Video Summarization Guided by Semantic Inverse Optimal Transport
mmfp2334	Adaptive Spatio-Temporal Directed Graph Neural Network for Parkinson's Detection using Vertical Ground Reaction Force

Paper id	Paper title
mmfp1375	TopicCAT: Unsupervised Topic-Guided Co-Attention Transformer for Extreme Multimodal Summarisation

All papers are presented also in the Poster Session W