

Introduction to the Issue on Emerging Techniques in 3-D

A PART from the conventional problems dealing with 3-D content, most of the current research efforts in 3-D involve more emerging techniques and focus on new aspects and issues, such as *modality*, *quality*, and *activity*. In other words, one of the most important emerging research efforts in 3-D area is based on fusion of conventional camera outputs with those captured by other modalities, such as active sensors, multi-spectral data or dynamic range images (in order to obtain better, cheaper and more reliable 3-D content). Another important area is devoted to the measurement and improvement of the quality of 3-D content, using still images or video, taking human visual system properties into account. A new paradigm, namely *Quality of Experience (QoE)* has been applied to 3-D content and become the main goal of many research efforts. Finally, 3-D information allows better segmentation and understanding of scene and actions.

This issue can be examined in four parts. The manuscripts in the first part are mainly devoted to fusion of different modalities with conventional camera outputs and there are four interesting papers in this “modality” section. The second part is about “quality” of 3-D content with four exciting manuscripts that provide new ways to measure the quality of 3-D videos or propose new methodologies to improve the quality of 3-D videos. The subsequent part is about analyzing the “activity” in the scene with another four stimulating contributions ranging from 3-D assisted segmentation of the scene, to 3-D representation of the objects in this scene and finally, to the analysis of their temporal activities. Lastly, the last section of the issue is dedicated to novel 3-D techniques, having three papers with contributions in robust extraction (interpolation) of 3-D point clouds, optimization of encoding latency in multi-view video and a new 3-D approach for high dynamic range imaging processing.

Our special issue starts with a novel solution to a fundamental problem in structured light systems, being described in “*Consistent Stereo-Assisted Absolute Phase Unwrapping Methods for Structured Light Systems.*” Phase unwrapping of the projected light by a projector to the scene is approached by a two-camera setup, allowing consistent solutions to this problem. Exploitation of phase consistency is achieved in either viewpoint or time, while both of these techniques are proven to improve the accuracy of reconstructed 3-D point clouds. In the next paper, “*Real-Time Distance-Dependent Mapping for a Hybrid ToF Multi-Camera Rig,*” a similar active sensor, a time-of-flight camera, is combined with an optical camera to yield real-time mapping of low-resolution depth measurements onto high-resolution color data. Real-time implementation is achieved elegantly by pixel associations that are described in a set of lookup tables, which solve the binocular disparity. Besides active sensors, another modality to fuse with conventional images could

be an infrared camera, as proposed in “*Multimodal Stereo Vision System: 3-D Data Extraction and Algorithm Evaluation.*” The matching between these two different modalities takes place only at sparse locations and is achieved with a gradient enriched mutual information metric. It is shown that reliable depth information can be extracted at these sparse points by such a color-IR combination. Finally, in “*Temporal-Dense Dynamic 3-D Reconstruction with Low Frame Rate Cameras,*” a dense low frame-rate camera rig is utilized to obtain a high frame-rate reconstruction by spatio-temporal fusion of the content. Although there is a single modality in the framework, the proposed solution depends upon fusion of spatio-temporal content by the help of shape context extracted with a dual-tree discrete wavelet transform.

Quality of 3-D content is another challenging topic. In this issue, there are a number of stimulating papers with interesting outcomes on this problem. The first paper, entitled “*Toward Assessing and Improving the Quality of Stereo Images,*” aims at fitting an objective model to subjective 3-D quality for stereo images. For this purpose, a number of features are proposed to assess 3-D quality of content and supervised learning is applied to determine a regression model to predict the 3-D quality of a stereo pair. In the subsequent paper, “*Edge-Based Reduced-Reference Quality Metric for 3-D Video Compression and Transmission,*” the same problem is extended to 3-D video (in 2-D+depth representation). Instead of an undesired full-reference quality metric, which requires the original content at the receiver side, the proposed technique is a reduced-reference method, in such a way that it only requires a binary edge-map of the original depth map to be transmitted for quality assessment. The simulations show that the proposed approach performs equivalent to its full-reference counterpart. In the following paper, “*Enhancement of Depth Maps with Alpha Channel Estimation for 3-D Video,*” rendering quality in color video and depth sensor scenario, is improved by a depth enhancement step together with a novel alpha-matting technique that yields more faithful blending of foreground and background objects during rendering, showing the effectiveness of combining depth and color alpha-matte in a linear fashion. Finally, for the 3-D systems that utilize depth image-based rendering for visualization, a hierarchical hole-filling technique is proposed in “*Hierarchical Hole-Filling For Depth-based View Synthesis in FTV and 3-D Video.*” In this manuscript, a fast and effective hole filling algorithm is proposed where the depth maps are not processed to avoid geometric distortions in the resulting 3-D video. The resulting quality outperforms the competing state-of-the-art algorithms.

Compared to conventional video, 3-D content brings extra clues about the scene so that makes its analysis more promising. For the analysis of the (scene) activity, the first step is typically segmentation, which is examined in “*Fusion of Geometry and*

Color Information for Scene Segmentation.” The key contribution of this manuscript is an automatic weighting procedure between color and depth, which is assumed to be captured by an active sensor. The automatic weighting is achieved by optimizing a metric that measures uniformity within regions, as well as irregularity between regions for both color and depth. Characterization of 3-D scene from its multi-view images is analyzed in “*Characterization of 3-D Volumetric Probabilistic Scenes for Object Recognition.*” The authors construct a volumetric probabilistic model of the scene from the observed image intensities and they are able to classify a number of objects (from aerial data) by using dense, as well as sparse, features obtained from the voxels by bag-of-words formulation. A survey on human 3-D pose and action recognition, entitled “*Human 3-D Pose Estimation and Activity Recognition from Multi-View Videos: Comparative Explorations of Recent Developments.*” follows next. This manuscript gives a thorough, quantitative and qualitative comparison between the state-of-the-art techniques on this topic. The last paper in scene activity part of the issue is entitled “*A Local 3-D Motion Descriptor for Multi-View Human Action Recognition from 4D Spatio-Temporal Interest Points.*” In this manuscript, in order to classify human actions from multi-view video, the spatio-temporal representation is achieved by a novel local 3-D descriptor, namely histogram of 3-D optical flow, while view-invariance is achieved through spherical harmonics. The authors demonstrate a clear improvement over similar techniques.

Aside from the 3-D emerging research efforts on modality, quality, and activity, there are further specific problems which we have considered as well. The remaining part of the special issue is devoted to three manuscripts, showcasing a few other representatives of promising directions. The first of these papers, entitled “*Noisy Depth Maps Fusion for Multi-view Stereo via Matrix Completion.*” applies a new technique, called *matrix completion*, to the problem of noisy depth map fusion, which is important during pair-wise depth extraction from multi-view content. To alleviate the effects of noise, a novel technique, namely log-sum penalty completion, is proposed with a non-convex objective function. Simulation results show a clear indication of superiority against state-of-the-art in performance-complexity tradeoff. Another interesting effort is related to optimizing encoding latency in multi-view compression, which is presented in the paper “*A Framework for the Analysis and Optimization of Encoding Latency for Multi-view Video.*” A new framework, *directed acyclic graph encoding latency* (DAGEL), is proposed to determine encoding latency for any encoding structure in multi-view coding. It is also possible to prune the structure until a target latency value is met by using this framework. This issue concludes with a manuscript dealing with a

novel problem in 3-D, with the title “*Rendering 3-D High Dynamic Range Images: Subjective Evaluation of Tone-Mapping Methods and Preferred 3-D Image Attributes.*” This paper considers the conversion of images captured in high dynamic range into conventional 8-bit low dynamic range 3-D displays. Supported by a number of subjective tests, it is concluded that there is clear distinction between global and local tone mapping operators, whereas all of them perform better with respect to low dynamic range images in 3-D.

We would like to thank the authors for submitting quality papers and our reviewers for their thoughtful and timely reviews. We also thank the Editor-in-Chief of this journal, Prof. Vikram Krishnamurthy, for his encouragement and support. The initiative and support from the IEEE Multimedia Technical Committee (MMTC) 3-D Rendering, Processing and Communications (3-DRPC) Interest Group to make this special issue successful is gratefully acknowledged. Finally, we are grateful to Ms. Rebecca Wollman for her assistance in assembling this issue.

A. AYDIN ALATAN, *Lead Guest Editor*
METU
TR-06800 Ankara, Turkey
alatan@eee.metu.edu.tr

JOERN OSTERMANN, *Guest Editor*
Leibniz Universität Hannover
D-30167 Hannover, Germany
ostermann@tnt.uni-hannover.de

LEVENT ONURAL, *Guest Editor*
Bilkent University
TR-06800 Ankara, Turkey
onural@bilkent.edu.tr

GHASSAN ALREGIB, *Guest Editor*
Georgia Institute of Technology
Atlanta, GA 30332 USA
alregib@gatech.edu

STEFANO MATTOCCIA, *Guest Editor*
University of Bologna
40136 Bologna, Italy
stefano.mattocchia@unibo.it

CHUNRONG YUAN, *Guest Editor*
University of Tuebingen
72076 Tuebingen, Germany
chunrong.yuan@uni-tuebingen.de