Introduction to the Issue on Visual Media Quality Assessment

THE increasing demand for digital image and video technologies, in applications as broad as entertainment and communications, security, monitoring, and medical imaging, has pushed to the forefront the need for accurate quality assessment strategies. Many factors can affect and/or impair the quality of visual media including acquisition, processing, compression, transmission, protection, display, printing, acquisition and reproduction systems. Visual media quality assessment aims at quantifying the quality of visual media, including still pictures, image sequences (video), 3-D visual data, and 3-D models, by means of quality metrics. These metrics vary with the considered applications, and range from metrics that measure specific visual impairments to those that assess the overall visual quality in the presence of various impairments. For applications and products that target human consumers, it is desirable to have metrics that will predict the perceived visual quality as measured with human subjects. Equally important are metrics that can measure a contextual visual quality in the presence or absence of impairments, and that can predict human perception as measured by the performance of a visual-based task.

Traditionally, visual quality assessment has been conducted using subjective tests in which human subjects are asked to rate the perceived visual quality of the displayed media according to a provided quality scale, with or without the presence of a reference visual medium, and based on specified criteria and conditions. A subjective quality metric can be computed by first assigning numerical scores to the individual ratings and then pooling the scores together in order to produce a single numerical score for each rated medium or test case. The pooling is commonly performed using averaging resulting in the Mean Opinion Score (MOS) subjective metric. When conducted properly, subjective tests result in subjective quality metrics that can very reliably predict the perceived visual quality. However, subjective quality metrics are costly, time-consuming and impractical as they cannot be integrated within real-world systems for real-time visual quality monitoring and control. These issues triggered the need to develop reliable objective quality metrics that can automatically assess the visual media quality as perceived by human observers. Automatic visual media quality assessment is crucial for efficiently monitoring and controlling the visual quality in deployed multimedia systems, and has the potential to impact next-generation systems by providing objective metrics for use during the design and testing stages and by reducing the need for extensive evaluation with human subjects.

Objective visual quality assessment metrics can be divided into full-reference, reduced-reference, and no-reference quality metrics. Full-reference visual quality metrics compare the to-be-assessed visual media to a reference, which is typically the original visual data. In many applications where the original visual data is not available, reduced-reference and no-reference metrics are used. Reduced-reference metrics make use of a set of reference features or characteristics, which could have been extracted from the original visual data. No-reference quality metrics attempt to predict the visual quality without any reference, which is very useful in practice but very challenging.

A great deal of interest and research have been devoted to the design and development of visual quality metrics, particularly full-reference and reduced-reference metrics for image quality assessment. However, for many applications, reliable automatic visual quality assessment is lacking, particularly those requiring no-reference visual quality assessment. In addition, there is a need for methods that can reliably assess the visual quality of video, high-definition visual content, color, multiview and 3-D visual media. There is also a need for novel application-specific visual media quality assessment methods that are tuned to specific visual tasks. For example, the ability to detect certain objects and structures would be an essential factor for assessing the quality of medical images. The reliability of visual quality metrics can be improved by augmenting those with multimodal quality assessments that take into account different modalities (e.g., auditory, affective, visual) in addition to context. Incorporating visual media quality assessment methods into various types of visual media processing applications is needed to build visually optimized algorithms and systems.

The goal of this special issue is to highlight the importance, challenges, and applications of visual media quality assessment and its interdisciplinary nature which includes vision science, optics, color science, signal processing, psychology, and biology. The issue contains twelve papers covering various aspects of visual media quality assessment. The first five papers deal with problems related to image (still picture) quality assessment, while the remaining seven papers are concerned with video quality assessment.

Among the image quality assessment papers, the first two papers present image quality metrics based on natural image statistics. Moorthy and Bovik address how to improve the full-reference SSIM quality index in "Visual Importance Pooling for Visual Quality Assessment" by taking into account the effect of "low quality regions" and visual importance on perceived errors. In "Reduced-Reference Image Quality Assessment Using Divisive Normalization-Based Image Representation," Li and Wang present a reduced-reference image quality assessment metric based on natural image statistics and a perceptually motivated divisive normalization transformation in the wavelet domain.

The next two image quality assessment papers address the topic of visual quality assessment in the context of image fusion applications. In "A Non-reference Image Fusion Metric Based on the Regional Importance Measure," Cvejic, Godsill, and Seppänen present a no-reference metric for the evaluation of image fusion algorithms by combining regional importance measurement and the visual information fidelity criterion. Building on the idea that, in several application scenarios, image fusion is used to facilitate specific visual-based tasks such as event mon-

Digital Object Identifier 10.1109/JSTSP.2009.2015485

itoring and object recognition by humans, Kaplan, Burks, Blum, Moore, and Nguyen deal with the evaluation of image quality in the context of specific tasks in "Analysis of Image Quality for Image Fusion via Monotonic Correlation." Monotonic analysis is introduced as a means to evaluate how well a set of full-reference and no-reference image quality features can predict human perception when measured in terms of object classification accuracy in fused images.

The fifth paper, by Li and Chen, deals with image quality assessment in the context of visual aesthetics. "Aesthetics Visual Quality Assessment of Paintings" deals with how to predict whether a painting is perceived to be very beautiful or not so beautiful by a human observer. While this problem appears highly subjective, the authors approach it as a machine learning problem using features representing both global and local characteristics of a painting.

The following seven papers address aspects in video quality assessment. Full reference video quality metrics that focus on exploiting the temporal information are presented in "Considering Temporal Variations of Spatial Visual Distortions in Video Quality Assessment" by Ninassi, Le Meur, Le Callet, and Barba, and in "Temporal Trajectory Aware Video Quality Measure" by Barkowsky, Bialkowski, Eskofier, Bitto, and Kaup.

Supra-threshold full-reference quality metrics for low bit rate video is the topic of "A Novel Video Quality Metric for Low Bit-rate Video Considering Both Coding and Packet-loss Artifacts" by Liu, Wang, Boyce, Yang and Wu. The investigation focuses on H.264 family codecs employed over wireless packet networks. Conducted subjective studies are used to develop a full-reference metric for quality degradation due to packet losses and compression.

In "Rule-Based No-Reference Video Quality Evaluation Using Additionally Coded Videos," Oelbaum, Keimel, and Diepold present a no-reference video quality metric to predict the perceived quality of AVC/H.264 compressed video.

The perceived quality of 3-D stereoscopic digital video is explored in "Quality Evaluation of Color Plus Depth Map-Based Stereoscopic Video" by Hewage, Worrall, Dogan, Villette, and Kondoz. The authors present the results of several subjective quality studies for stereoscopic video sequences.

Drelie Gelasca and Ebrahimi address how to assess the quality of video object segmentation in "On Evaluating Video Object Segmentation Quality: A Perceptually Driven Objective Metric." The authors present a full-reference objective metric for video segmentation quality evaluation. The authors also provide both a subjective and objective performance comparison of existing video object segmentation systems for several video-based applications.

In "Consumer Opinions About Frequency Of Artifacts In Digital Video," Cermak presents the results of two studies on consumer opinions about artifacts in digital video. In the first study, consumers were asked to report on how often they perceived various classes of artifacts in broadcast video. In the second, a different set of consumers were asked to place a cost on how much they were willing to pay to not have certain artifacts appear. These studies provide insight into how the actual end users of digital video systems rank various types of artifacts.

LINA J. KARAM, *Lead Guest Editor*
Arizona State University
Tempe, AZ 85287–5706 USA

TOURADJ EBRAHIMI, *Guest Editor*
Q2S-NTNU
Trondheim, 7034 Norway
EPFL-STI-ITS-LTS1
Lausanne, CH 1015 Switzerland

SHEILA S. HEMAMI, *Guest Editor*
Cornell University
Ithaca, NY 14853 USA

THRASYVOULOS N. PAPPAS, *Guest Editor*
Northwestern University
Evanston, IL 60208–3118 USA

ROBERT J. SAFRANEK, *Guest Editor*
Benevue, Inc.
Warren, NJ 07059 USA

ZHOU WANG, *Guest Editor*
University of Waterloo
Waterloo, ON N2L 3G1 Canada

ANDREW B. WATSON, *Guest Editor*
NASA Ames Research Center
Moffett Field, CA 94035–1000 USA

Lina J. Karam (S'91–M'95–SM'03) received the B.E. degree in computer and communications engineering from the American University of Beirut, Beirut, Lebanon, in 1989, and the M.S. and Ph.D. degrees in electrical engineering from the Georgia Institute of Technology, Atlanta, in 1992 and 1995, respectively.

She is currently an Associate Professor in the Electrical Engineering Department, Arizona State University, Tempe. Her research interests are in the areas of image and video processing, image and video coding, visual quality assessment, human perception, error-resilient source coding, medical imaging, and digital filtering. She worked at Schlumberger Well Services on problems related to data modeling and visualization, and in the Signal Processing Department of AT&T Bell Labs on problems in video coding during 1992 and 1994, respectively.

Prof. Karam is the recipient of an NSF CAREER Award. She served as the Chair of the IEEE Communications and Signal Processing Chapters in Phoenix in 1997 and 1998. She also served as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING from 1999 to 2003 and of the IEEE SIGNAL PROCESSING LETTERS from 2004 to 2006, and as a member of the IEEE Signal Processing Society's Conference

Board from 2003 to 2005. She is currently serving as an Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING and as the Technical Program Co-Chair of the 2009 IEEE International Conference on Image Processing. She also serves on the editorial board of the *Foundations and Trends in Signal Processing* journal. She is an elected member of the IEEE Circuits and Systems Society's DSP Technical Committee and of the IEEE Signal Processing Society's IVMSP Technical Committee. She is a member of the Signal Processing, and Circuits and Systems societies of the IEEE.

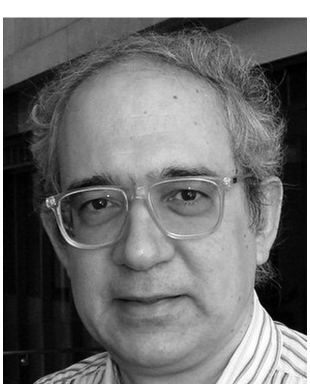
Touradj Ebrahimi is currently a Professor at EPFL, Lausanne, Switzerland, heading its Multimedia Signal Processing Group. He also serves as an Adjunct Professor with the Center of Quantifiable Quality of Service at Norwegian University of Science and Technology (NTNU). His research interests include still, moving, and 3-D image processing and coding, and visual information security. He is the author or the coauthor of more than 200 research publications and holds 14 patents.

Prof. Ebrahimi has been the recipient of various distinctions and awards, such as the IEEE and Swiss national ASE award, the SNF-PROFILE grant for advanced researchers, Four ISO-Certificates for key contributions to MPEG-4 and JPEG 2000, and the Best Paper Award of IEEE TRANSACTIONS ON CONSUMER ELECTRONICS. He became a Fellow of the international society for optical engineering (SPIE) in 2003. He has initiated more than two dozen National, European, and International cooperation projects with leading companies and research institutes around the world. He is also the head of the Swiss delegation to MPEG, JPEG, and SC29, and acts as the Chairman of Advisory Group on Management in SC29. He is a member of Scientific Advisory Board of various start-up and established companies in the general field of information technology. He has served as Scientific Expert and Evaluator for research funding agencies such as those of European Commission, the Greek Ministry of Development, the Austrian National Foundation for Scientific Research, the Portuguese Science Foundation, as well as a number of venture capital companies active in the field of information technologies and communication systems.

Sheila S. Hemami (S'89–M'95–SM'03–F'08) received the B.S. degree (summa cum laude) in electrical engineering from the University of Michigan, Ann Arbor, in 1990, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1992 and 1994, respectively.

She was with Hewlett-Packard Laboratories, Palo Alto, CA, in 1994. In 1995, she joined the faculty of the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY, where she is currently a Professor and Director of the Visual Communications Lab. She has held visiting positions at Ecole Polytechnique Federale de Lausanne (EPFL WISH Distinguished Visiting Professor), Princeton University, Princeton, NJ, and Rice University, Houston, TX (TI Distinguished Visiting Professor), and in 2001, she visited the Faculte de Sciences, Rabat, Morocco, as a Fulbright Distinguished Lecturer.

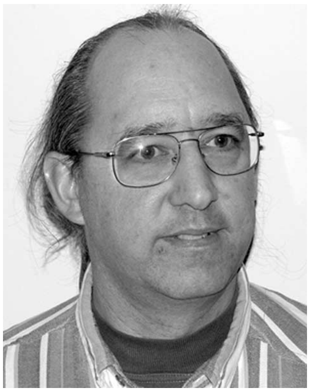
Dr. Hemami is currently the Editor-in-Chief for the IEEE TRANSACTIONS ON MULTIMEDIA and serves as an elected member of the Signal Processing Society Board of Governors. She chaired the IEEE Image and Multidimensional Signal Processing Technical Committee, and has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING. She has served on various program committees and organizing committees. In 1997, she received a National Science Foundation CAREER Award. She held the Kodak Term Professorship of Electrical Engineering at Cornell University from 1996 to 1999. In 2000, she received the Eta Kappa Nu C. Holmes MacDonald Outstanding Teaching Award (a national award), and she has won numerous teaching awards at Cornell University. She was a finalist for the Eta Kappa Nu Outstanding Young Electrical Engineer in 2003. In 2005, she received the Alice H. Cook and Constance E. Cook Award at Cornell University for her leadership of the Women in Science and Engineering committee. Dr. Hemami is a member of Eta Kappa Nu and Tau Beta Pi.



Thrasyvoulos N. Pappas (M'87–SM'95–F'06) received the B.S., M.S., and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, in 1979, 1982, and 1987, respectively.

From 1987 to 1999, he was a Member of the Technical Staff at Bell Laboratories, Murray Hill, NJ. In 1999, he joined the Department of Electrical and Computer Engineering (now EECS), Northwestern University, Evanston, IL, as an Associate Professor. His research interests are in image and video quality and compression, perceptual models for image processing, image and video analysis, model-based halftoning, and multimedia signal processing.

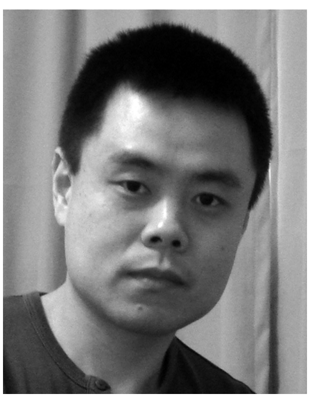
Dr. Pappas has served as an elected member of the Board of Governors of the Signal Processing Society of IEEE (2004–2007), chair of the IEEE Image and Multidimensional Signal Processing Technical Committee, Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, and technical program Co-Chair of ICIP 2001, the Symposium on Information Processing in Sensor Networks (IPSN 2004), and ICIP 2009. He is a Fellow of SPIE. Since 1997, he has been Co-Chair of the SPIE/IS&T Conference on Human Vision and Electronic Imaging. He has also served as Co-Chair of the 2005 SPIE/IS&T Electronic Imaging Symposium.



Robert J. Safranek (SM'89) received the B.S.E.E., M.S.E.E., and Ph.D. degrees from Purdue University, West Lafayette, IN, in 1980, 1982, and 1986, respectively.

From 1986 to 2000, he was with the Signal Processing Research Department, Bell Laboratories. While at Bell Labs, he worked on developing visual models for image and video quality evaluation and compression. He was also a member of the team that helped develop the U.S. HDTV standard. Since 2001, he has held the position of CTO/founder of Benevue, Inc., Warren, NJ, where he has led the development of their color formulation and matching products for the paint, dye, and printing industries. He holds over two dozen patents and has published extensively. His current technical interests include visual models for quality assessment and compression, color matching, and multimedia information retrieval.

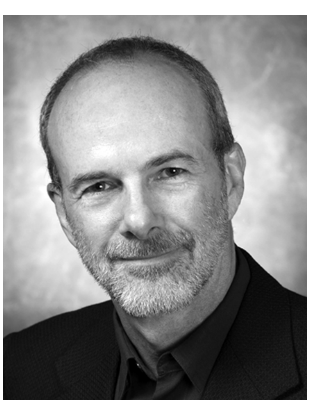
Dr. Safranek was awarded the 1995 IEEE Donald G. Fink Prize, was co-program chair for the 2007 ICIP conference, and is a former member of the IEEE Image and Multidimensional Signal Processing Technical Committee.



Zhou Wang (S'99–A'01–M'02) received the Ph.D. degree from The University of Texas at Austin in 2001.

He is currently an Assistant Professor in the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. Before that, he was an Assistant Professor in the Department of Electrical Engineering, The University of Texas at Arlington, a Research Associate at Howard Hughes Medical Institute and New York University, and a Research Engineer at AutoQuant Imaging, Inc. His research interests include image processing, coding, communication, and quality assessment; computational vision and pattern analysis; multimedia coding and communications; and biomedical signal processing. He has more than 60 publications and one U.S. patent in these fields and is an author of *Modern Image Quality Assessment* (Morgan & Claypool, 2006).

Dr. Wang is an Associate Editor of the IEEE SIGNAL PROCESSING LETTERS and of *Pattern Recognition*.



Andrew B. Watson did undergraduate work at Columbia University, New York, and received the Ph.D. degree in psychology from the University of Pennsylvania, Philadelphia, in 1976.

He subsequently held postdoctoral positions at the University of Cambridge, Cambridge, U.K., and at Stanford University, Stanford, CA. Since 1980, he has been with the NASA Ames Research Center, Moffett Field, CA, where he is currently a Senior Scientist for Vision Research, and where he works on models of vision and their application to visual technology. He has developed widely used models of motion perception and spatial pattern perception, as well as methods of efficient data collection. He is the author of over 100 papers on topics such as spatial and temporal sensitivity, motion perception, image quality, and neural models of visual coding and processing. He is the author of five patents, in areas such as image compression, video quality, and detection of artifacts in display manufacturing. In 2001, he founded the *Journal of Vision* (http://journalofvision.org) where he now serves as Editor-in-Chief. He also serves as an Associate Editor for the journal *Displays*.

Dr. Watson is a Fellow of the Optical Society of America and also serves as the Vice Chair for Vision Science and Human Factors of the International Committee on Display Measurement. In 1990, he received NASA's H. Julian Allen Award for outstanding scientific paper, and in 1993 he was appointed Ames Associate Fellow for exceptional scientific achievement. He is the 2007 recipient of the Otto Schade Award from the Society for Information Display and the 2008 winner of the Special Recognition Award from the Association for Research in Vision and Ophthalmology.