

WestminsterResearch

<http://www.westminster.ac.uk/research/westminsterresearch>

Special Section Guest Editorial: Image/Video Quality and System Performance

**Mohamed-Chaker Larabi
Sophie Triantaphillidou
Andrew B. Watson**

This is a copy of the final published version of an article published in Journal of Electronic Imaging, Volume 23 (6) (2014), pp. 061101-1 - 061101-2.

© 2014 by the authors, SPIE and IS&T. This is an open access article distributed under the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The published version is available at: <http://dx.doi.org/10.17645/mac.v3i2.207>

The WestminsterResearch online digital archive at the University of Westminster aims to make the research output of the University available to a wider audience. Copyright and Moral Rights remain with the authors and/or copyright owners.

Whilst further distribution of specific materials from within this archive is forbidden, you may freely distribute the URL of WestminsterResearch: (<http://westminsterresearch.wmin.ac.uk/>).

In case of abuse or copyright appearing without permission e-mail repository@westminster.ac.uk

Journal of Electronic Imaging

JElectronicImaging.org

Image/Video Quality and System Performance

Mohamed-Chaker Larabi
Sophie Triantaphillidou
Andrew B. Watson

Image/Video Quality and System Performance

Mohamed-Chaker Larabi

Université de Poitiers
XLIM Signal Image and Communication Department
Bvd. Marie Et Pierre Curie
BP 30179, 86962 France
E-mail: chaker.larabi@univ-poitiers.fr

Sophie Triantaphillidou

University of Westminster
Faculty of Media, Arts and Design
Watford Road, Harrow
HA1 3TP, United Kingdom
E-mail: triants@westminster.ac.uk

Andrew B. Watson

NASA Ames Research Center
MS 262-2
Moffett Field, California 94035-1000, United States
E-mail: andrew.b.watson@nasa.gov

Rapid developments in display technologies, digital printing, imaging sensors, image processing and image transmission are providing new possibilities for creating and conveying visual content. In an age in which images and video are ubiquitous and where mobile, satellite, and three-dimensional (3-D) imaging have become ordinary experiences, quantification of the performance of modern imaging systems requires appropriate approaches. At the end of the imaging chain, a human observer must decide whether images and video are of a satisfactory visual quality. Hence the measurement and modeling of perceived image quality is of crucial importance, not only in visual arts and commercial applications but also in scientific and entertainment environments. Advances in our understanding of the human visual system offer new possibilities for creating visually superior imaging systems and promise more accurate modeling of image quality. As a result, there is a profusion of new research on imaging performance and perceived quality.

During the last decade important research efforts have been put to understand and model human visual attention. This yielded several computational saliency models that try to find the most attractive objects based on image/video characteristics. Most existing models are low-level feature-oriented (bottom-up) approaches; only a few are or task-oriented (top-down). The interest in such models is due to the potential optimization that can be obtained when the salient areas are known. This likely helps preserve the quality of some of the scene objects when applying processing like compression, watermarking, transmission, and so on. Moreover, the performance of image processing algorithms and the quality of experience linked to them, are fundamentally associated with the salient information. Impairments falling in the salient areas will have a disproportionate impact on the quality.

Quality assessment has been a challenge in the last decade and many issues are still open in this field. An impressive number of full reference metrics dedicated to images have been developed. Many of them are well correlated to human judgment in the framework of multimedia applications. However, only a few metrics have attempted blind (no reference) or semi-blind (reduced reference) image quality

assessment. First, because these type of metrics are highly dependent upon the impairment type. Second, it is always difficult to select side information that is both reduced and efficient in representing the original image. These problems increase when dealing with video content, where the temporal variation of both content and quality makes the scope very different from the image one. Furthermore, with the advent of new technologies like 3-D or multiview, quality is facing new challenges since the involvement of the visual perception is increased. Moreover, depth is not only an additional dimension but is also a feature changing the full interpretation of a scene. Therefore, the research community has to improve its understanding of the binocular vision mechanisms such as the binocular rivalry, suppression and fusion, in order to be able to predict the perceptual impact of the various impairments.

High dynamic range (HDR) is expected to be the next stage of multimedia imaging. It allows avoidance of the problems caused by tone mapping operators and other grayscale processing. However, as a new technology, little is known about how to assess the quality of HDR images and video. An important research effort is now under way to understand HDR and to develop tools and quality metrics that fulfill the challenges of the quality of experience.

This special section presents some of the recent advances in image and video quality assessment and system performance with applications to multimedia, consumer camera, image and video coding, and other related topics. The papers of this special section can be grouped into four categories, according to the addressed subject.

The first category addresses image and video quality metrics. First, a psychometric study evaluates a number of no-reference image quality metrics, based on natural scene statistics, which were previously shown to predict successfully the subjective quality of digital image processing algorithms. The best performing metrics, BIQI and NIQE, are shown to perform equally well on two tests with four printed images and different graded papers, while NIQE is also robust to scene content variations. In the second study concerned with the measurement of image quality, a no-reference image quality metric uses image features derived from the image's contourlet transform to obtain structural image

information and a singular-value decomposition is proposed. The advantage of the metric is that it does not require training or learning and is shown to deliver predictions that correlate well with human observations. Finally, multimodal quality assessment is very important since it corresponds to the real-life experience with visual content. Despite the relevance of this field, few approaches have been proposed in the literature. One contribution of this special section addresses this problem by running psychophysical experiments to clarify the interaction between audio and video. Based on their subjective findings, authors propose objective metrics combining audio and video quality in various ways.

Three papers in this issue consider visual saliency and attention. In the first, the problem of saliency detection is approached by selecting image features from multiple image views with different contributions. A multimanifold ranking algorithm determining final saliency is presented and tested using qualitative and quantitative comparisons and different image databases. In the second, a number of full-reference video quality metrics were assessed with and without the incorporation of visual saliency models. It is shown that, overall, computational models of attention benefit all selected metrics, but predominantly those designed to predict spatial degradations. Finally, authors of the third paper explored two different visual tasks related to quality, i.e. quality estimation and difference estimation. Eye-tracking results showed a different strategy for each of the tasks. Visual characteristics (e.g. fixations) are different in position, duration, and role, leading to the conclusion that instructions may have a noticeable influence on the strategy used.

Image and video coding is the focus of another set of papers. In the first, authors propose an optimized parallel implementation of the DCT transform using OpenCL (open computing language). The results show interesting speed-up factors in heterogeneous environments. Authors of the second paper propose a Wyner-Ziv video coding scheme with low bit-rate fluctuations based on a symmetric coding structure. The bit-rate stability is seen as a great advantage for many applications with constrained communication bandwidth. Recently, a joint effort between ITU and ISO allowed the development of a new video coding standard called HEVC (high efficiency video coding) based on a quad-tree structure to provide variable transform sizes in the transform coding process. The third paper proposes a similarity-check scheme to efficiently reduce the transform unit (TU) candidate modes. The proposed optimization allowed an important saving of the TU encoding time for almost invisible impairments. Finally, the effect of coding and transmission is evaluated on the performance of a state-of-the-art pedestrian detection algorithm. "Quality aware" spatial image statistics are used to blindly categorize distorted video frames by distortion type and level in a blind way.

In a paper concerned with feature subsets that predict image quality of consumer cameras, the authors have used the new CID2013 large database with consumer camera photographs, to classify images using principal component analysis. Images were classified in terms of sharpness and noise

energy, with the sharpness dimension including lightness, resolution and contrast. A feature subset model is proposed that successfully predicts human observations. The second paper proposes a benchmark to solve the problem of how to combine quality and speed metrics for mobile phone camera evaluation. Vergence-accommodation conflict has been studied for 3-D in order to understand the accumulation of visual fatigue. A paper in this special issue has been dedicated to investigate this conflict for holographic stereograms. It showed that improved full-parallax holographic stereogram can control the focusing depths of points and guarantee consistency between the vergence and accommodation distances.

Several people deserve to be acknowledged for the success of this special section. We would like to thank authors for submitting their high-quality papers, and reviewers for dedicating their time and expertise to selecting the best papers for publication. Special thanks go to Editor-in-Chief Dr. Gaurav Sharma and the editorial staff for their full support throughout the process of this special section.

Mohamed-Chaker Larabi received his PhD from the University of Poitiers (2002). He is currently the associate professor in charge of the perception, color, and quality activity at the same university. His actual scientific interests deal with image and video coding and optimization, 2-D and 3-D image and video quality assessment, and user experience. He published over 100 papers in computer vision, image quality, and human perception. He acted as chair/co-chair of the program committee of several conferences and guest editor for special issues. He is a member of MPEG and JPEG committee (since 2000), and chair of the Advanced Image Coding Group and the Test and Quality Subgroup. He serves as a member of divisions 1 and 8 of CIE, is a member of IS&T, and a senior member of IEEE.

Sophie Triantaphillidou is an associate professor in imaging science and the director of the Imaging Technology Research Group of the University of Westminster, UK. She received her PhD from the University of Westminster in imaging science in 2002. She has published over 40 papers relating to imaging system performance, image quality, and human vision and is the co-editor and a main contributor of the 10th edition of *Manual of Photography*, published by Focal Press in 2010. She has been the co-chair of the SPIE/IS&T Electronic Imaging: Image Quality and System Performance conference since 2012 and has been involved in several national and international program committees relating to imaging. She was the recipient of the Davis Medal of the Royal Photographic Society in 2012, awarded for her contributions to digital imaging science.

Andrew B. Watson is the Senior Scientist for Vision Research at NASA Ames Research Center in California. He was an undergraduate at Columbia University, received his PhD in psychology from the University of Pennsylvania, and did postdoctoral work at the University of Cambridge in England. He is the author of over 100 papers and seven patents on topics in vision science and imaging technology. In 2001, he founded the *Journal of Vision* (<http://journalofvision.org>), where he served as editor-in-chief. He is a fellow of the Optical Society of America, of the Association for Research in Vision and Ophthalmology, and of the Society for Information Display. He is vice chair for vision science and human factors of the International Committee on Display Measurement.