Preprint of paper published in Signal Processing: Image Communication 2015

Signal Processing: Image Communication 30 (2015) 000–000

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

Guest Editorial

Recent Advances in Vision Modeling for Image and Video Processing

Computational models of the human visual system (HVS) play an instrumental role in the design and operation of numerous image and video processing algorithms. HVS models have proved beneficial not only for boosting an algorithm's performance, but also for improving its reliability and robustness to changing conditions.

Many processing algorithms operate on images and videos that are meant to be viewed by humans. For such algorithms, HVS models can help on multiple fronts. First, many HVS models can be directly incorporated into one or more stages of a processing algorithm. This approach is quite beneficial because it allows an algorithm to take into account aspects of visual perception that would otherwise be neglected.

Second, even if an HVS model is too time-consuming or resource-intensive to be integrated into an algorithm, the model can still be useful for guiding an algorithm's design. For example, an HVS model can be used to evaluate the perceptual impacts of a particular processing stage, which then allows the engineer to make more perceptually optimal design decisions, and thereby allows the engineer to indirectly incorporate aspects of visual perception into an algorithm.

A third benefit of HVS modeling is the possibility to learn from the processing strategies employed in the models. Many effective analysis techniques such as sparse coding, divisive normalization, and adaptive feature selection have proven crucial for accurately modeling psychophysical results. Subsequently, these approaches have found widespread use in a variety of algorithms such as coding, segmentation, and quality assessment. In this way, an HVS model can serve as an intermediary for transferring biological processing strategies employed by the HVS into computational processing strategies employed by an algorithm.

This special issue represents a diverse set of research efforts on recent advances in HVS modeling for image and video processing—including psychophysical experiments, improved perceptual and saliency modeling, quality assessment, and novel applications of HVS models. Ultimately, all of the studies promote a better understanding of the theory and application of HVS models for image, video, and multimedia applications.

Psychophysics and Perceptual Models: The first set of papers deals with new psychophysical studies and perceptual models.

In their paper titled "*Scene Masking is Affected by Trial Blank-Screen Luminance*," Freeman, Loschky, and Hansen describe a psychophysical study designed to investigate the influence of inter-screen luminance contrast (ISLC) on scene masking. Blank screens are commonly used in psychophysical experiments between image presentations, thus causing ISLC between the blank screen and the image. The results of their study demonstrate that gray blank screens vs. black/white blank screens induce lesser extraneous ISLC masking, depending on the stimulus onset asynchrony.

In their paper titled "*Motion Silencing of Flicker Distortions on Naturalistic Videos*," Choi, Cormack, and Bovik provide both a psychophysical study on and associated model for the ability of motion to silence the visibility of flicker distortions. Flicker visibility is found to be strongly reduced when the speed of coherent motion is large, and the effect is pronounced when the video quality is poor. A spatiotemporal backward masking model is proposed based on these findings and flicker visibility is predicted based on a learned model of neural flicker adaptation processes. Experimental results show that the proposed model correlates well with human perception of flicker distortions.

In their paper titled "*What Can We Expect from a Classical V1-MT Feedforward Architecture for Optical Flow Estimation?*" Solari, Chessa, Medathati, and Kornprobst describe a joint V1 and MT model for optical-flow estimation. An essential contribution of this paper is to show how a neural model can be enriched to deal with real sequences and to provide a baseline for future development of bio-inspired and scalable computer vision algorithms.

Video Compression: The next set of papers deals with subjective evaluation and HVS modeling for video coding.

In their paper titled "*Color Difference Weighted Adaptive Residual Preprocessing Using Perceptual Modeling for Video Compression*," Shaw, Allebach, and Delp, provide a new contrast sensitivity model that takes into account color perception for use in video coding. Their model guides an associated perceptual strategy to selectively attenuate inter-frame differences. The approach is shown to yield significantly improved inter-frame compression, and is applicable to any standardized video coding scheme.

In their paper titled "*Subjective Evaluation of Super Multi-View Compressed Contents on High-End Light-Field 3D Displays*," Dricot, Jung, Cagnazzo, Pesquet, Dufaux, Kovács, and Adhikarla present a series of subjective experiments and analyses on the impact of compression on super multi-view (SMV) 3D video. Several coding configurations for SMV content and inter-view prediction structures are compared. The use of SMV is concluded to be realistic according to next generation compression technology requirements, but an improvement of compression efficiency, depth estimation and view synthesis is recommended.

Visual Saliency: The next set of papers deals with new studies and models of visual saliency in images and videos.

In their paper titled "*Perceived Interest and Overt Visual Attention in Natural Images*," Engelke and Le Callet describe a study designed to provide a deeper understanding of the relationship between perceived interest and overt visual attention. Towards this goal, a dedicated region-of-interest selection experiment and an eye gaze tracking experiment were performed. The results reveal that there is indeed a strong relationship between perceived interest and overt visual attention for a wide range of natural scenes.

In their paper titled "*Temporal Resolution vs. Visual Saliency in Video: Analysis of Gaze Patterns and Evaluation of Saliency Models*," Cheon and Lee describe a study designed to investigate the effects of variable frame-rate video on visual saliency. Via an eye-tracking experiment, they demonstrate that both the average eye-path and subject-wise variability of the eye-path are influenced by variations in the frame-rate. They also demonstrate that state-of-the-art saliency models cannot simultaneously achieve both accuracy and robustness in predicting the paths for variable frame-rate video.

In their paper titled "*Goal-Oriented Top-Down Probabilistic Visual Attention Model for Recognition of Manipulated Objects in Egocentric* Videos," Buso, Gonzalez-Diaz, and Benois-Pineau describe a new top-down probabilistic saliency model for egocentric videos. The model is probabilistically defined using both global and local appearance features extracted from automatically segmented arm areas and objects. A psycho-visual experiment has been conducted in a guided framework and the obtained results show that the proposed approach outperforms several popular bottom-up saliency approaches on a well-known egocentric dataset.

Quality Assessment: The next set of papers deals with new studies and algorithms for quality assessment and QoE assessment.

In their paper titled "*Quality of Experience of Adaptive Video Streaming: Investigation in Service Parameters and the Subjective Quality Assessment Methodology*," Tavakoli, Brunnström, Gutiérrez, and Garcia present two studies regarding QoE of adaptively streamed video. The first study compares two different testing methodologies for obtaining subjective QoE scores for HAS video. The second study investigates the impacts of various HAS parameters on QoE. The authors present several important findings that can help guide QoE optimizations for adaptive streaming services.

In their paper titled "*Image Retargeting Quality Assessment Based on Support Vector Regression*," Liu, Lin, Chen, and Zhang present a QA scheme for retargeted images based on four quality factors and support vector regression. Their quality factors separately and independently capture both shape distortions and visual content changes, and these measures are then fused to estimate the overall quality. Testing on several databases verifies the improvements of their approach over current techniques.

In their paper titled "*Lightweight Implementation of No-Reference (NR) Perceptual Quality Assessment of H.264/AVC Compression*," Leszczuk, Kowalczyk, Janowski and Papir propose a lightweight no-reference metric to objectively assess H.264/AVC video quality in real-time. This model takes into account typical artefacts introduced by hybrid block-based motion-compensated predictive video codecs such as blocking and temporal flickering. Experimental evaluations show that the proposed metric performs better than popular full reference metrics such as the structural similarity index.

Acknowledgements

The guest editors would like to thank the authors and anonymous reviewers for their great efforts in making this issue possible. A special thanks to the Editor-in-Chief, Prof. Frederic Dufaux, and to the editorial and production staff at SPIC (Wen Latie and Pattan Salma Begum) for their extraordinary support of this special issue.

D. M. Chandler would like to acknowledge US NSF Grant 1054612 for support of this research theme.

Damon M. Chandler (Guest Editors)
Department of Electrical and Electronic Engineering, Shizuoka University, Hamamatsu, Japan

Ulrich Engelke
Commonwealth Scientific and Industrial Research Organization (CSIRO), Hobart, Australia

Yuukou Horita
Department of Electronics and Computer, University of Toyama, Toyama, Japan

Kalpana Seshadrinathan
Intel Labs, Santa Clara, CA, USA