EDITORIAL

Special issue on human vision and information theory

Vittoria Bruni · Domenico Vitulano · Zhou Wang

1 Introduction

In the last two decades, there has been an increasing interest in the signal and image processing community toward models and techniques based on Human Vision. These models and techniques have been employed in many application fields, including image and video quality assessment, image and video compression, image restoration and enhancement, image recognition and classification, computer graphics and robotics—just to name a few. Human Visual System (HVS)-based approaches are highly desirable because of two reasons:

– it allows to design algorithms for objective evaluation of quality in a way that is consistent with subjective human evaluation;
– it enables automatic algorithms adapted for optimal perceptual quality.

Nevertheless, despite a large variety of interesting empirical results, theoretical principles that are able to promote further development of efficient and reliable models are still lacking. Presently, an exciting new research trend is to exploit

V. Bruni (✉)
Department of SBAI, Faculty of Engineering, University of Rome "Sapienza", Via A. Scarpa 16, 00161 Rome, Italy
e-mail: vittoria.bruni@sbai.uniroma1.it

D. Vitulano
Istituto per le Applicazioni del Calcolo "M. Picone", C.N.R.,
Via dei Taurini 19, 00185 Rome, Italy
e-mail: d.vitulano@iac.cnr.it

Z. Wang
Department of Electrical and Computer Engineering, University of Waterloo, 200 University Ave. W., Waterloo, ON N2L 3G1, Canada
e-mail: Z.Wang@ece.uwaterloo.ca

well-established results of Information Theory to model the information processing mechanisms in the HVS, so as to design perceptually tuned image processing algorithms and systems. This strategy provides a new perspective on image content where image pixels are no longer seen as simple probabilistic data but should be accounted for by the HVS limits. As a result, HVS-based methods may be used not only to exploit but also to refine some classical (upper and lower) bounds of Information Theory.

The goal of the present Special Issue is to address the latest developments in the area of Human Vision and Information Theory, and their applications in the field of image and video processing.

This Issue starts with a comprehensive survey on information-theoretic approaches oriented to solving a variety of visual quality assessment (QA) problems. All three categories of image QA models are treated for both images and videos: full reference (FR), reduced reference (RR) and no reference (NR) measures—with focuses on the contribution of Information Theory in each of them. Specifically, mutual information and conditional Kolmogorov complexities are incorporated into FR measures; relative entropy or entropic differences are used in RR measures, and Renyi entropy and mutual information–oriented weighting strategies are employed in NR QA algorithms.

QA measures are further investigated in the second contribution of the Issue, where the Normalized Perceptual Information Distance (NPID) measure is proposed. Starting from the theory of Kolmogorov complexity and exploiting the properties of the wavelet transform, maximum likelihood estimation and least square regression are combined for the design of an effective distortion measure that well correlates with HVS.

Some theoretical links between the Jensen–Shannon divergence (JSD) and HVS features are investigated in the

third contribution. It is proved that JSD can be formally written in terms of the well-known structural similarity measure (SSIM), some powers of the Weber's law and an additional component tied to the skewness of the distribution of the involved data. This leads JSD to have a trend similar to other popular Quality Assessment (QA) measures in the presence of common degradation types.

However, HVS is not a perfect device and not all QA measures perfectly fit its behavior in different conditions: high and low quality. This is the topic of the fourth contribution of this Issue. Its main idea consists of combining different quality metrics, including SSIM and VIF, in order to further improve their performance.

Despite the limits of HVS, it often works better than manmade acquisition devices, which often have serious problems in critical acquisition conditions. This aspect, embedded in one of the fundamental problems of Information Theory (data compression), is the topic of the fifth contribution of this Issue. In particular, the potential of embedding HVS in a compression framework is exploited in two approaches designed to adaptively (JPEG) coding color images inside digital cameras. The main peculiarity of both approaches is that the achieved compression, although lossy, introduces only invisible distortions. Both approaches are shown to improve average compression ratio (CR) by more than two times as compared to Super-High Quality (SHQ) mode widely used in consumer digital cameras.

The study on QA-related issues ends with the sixth contribution, which is concerned about the detection of blocking artifacts in video frames. A measure able to evaluate how much the video is degraded by blocking artifacts is proposed. Information-Theoretic measures are integrated with models of Human Perception, to account for the visibility of the artifact on different image contents. The resulting method is suitable to operate on the decoded version of the frames at the final stage of the video processing chain and it also shows good correlations with subjective scores.

Joint Information theory and HVS models can also help better understand and render real-world objects. This is the topic of the seventh and eighth contribution of this Issue. The former analyzes the information associated with an object in a visibility channel between a set of viewpoints and the polygons of the object itself. This approach allows to quantify object information as well as to select the N best views for exploring it. The latter paper proposes an image-based simplification method for textured triangle meshes that preserves the structural appearance of textured models. Again, an information channel created between a sphere of viewpoints and the object texture regions is employed. Such a channel enables to define both the Shannon entropy and the mutual information associated with each viewpoint and their corresponding generalizations based on Harvda-Charvat-Tsallis entropy.

The last two contributions of this Issue regard HVS as the result of the Darwin evolution, that is responsible for the quick human perception of many phenomena and events.

In particular, the ninth contribution focuses on a specific problem in Surveillance: the influence of image distortion on face classification techniques. This limit hinders their use in large scale commercial applications, which are typically run in uncontrolled settings. The problem can be addressed using the concept of "feature representativeness" i.e., a subject gallery that includes a sufficient amount of possible variations, in order to allow a correct recognition in different situations. The combination of Information Theory tools with HVS gives, again, an effective solution to the problem, as shown in the ninth contribution of the Issue, which investigates the role of the entropy computed over a set of samples of the same subject.

The tenth and last contribution of the Issue deals with the automatic shot boundary detection and keyframe selection in video. Two different information-theoretic approaches are proposed. They quantify the similarity between two frames by means of Tsallis mutual information and Jensen–Tsallis divergence, respectively. It is shown that Tsallis mutual information for the HSV and Lab color spaces with only 8 regular bins for each color component substantially outperforms previously proposed methods based on mutual information and Jensen–Shannon divergence.

We hope that the contributions in this Special Issue demonstrate the excitement arised by the marriage between HVS and Information Theory, which is opening up new research opportunities in Information Theory and is producing fruitful novel and effective techniques in Image and Video Processing.