

Perception, Aesthetics, and Emotion in Multimedia Quality Modeling

Introduction

The volume of multimedia data we handle on a daily basis is growing exponentially due to the availability of ubiquitous and cheap sensors, sharing platforms and new social trends. Artificial intelligence techniques have proven useful for interpreting this preponderance of data. In the last decades, many quality models have been proposed by mimicking the process of human perceiving multimedia data. Such perceptual quality models can benefit a rich variety of multimedia applications. For example, an effective photo aesthetics prediction module can help photographers crop an aesthetically pleasing sub-region from an original poorly framed photo. Also, a successful photo management system can rank videos based on the human perception of video quality (i.e., frame aesthetics, stability, and coherence), thereby the users can conveniently select their favorite pictures into albums. Last, different criteria have been developed to select visual or acoustic features for various multimedia applications, e.g., multimodal event detection, real-time speech recognition, and cross-media retrieval.

Extensive research efforts have been dedicated to designing perceptual quality models, but effective tools to manipulate the quality prediction are still at their infancy. As far as we know, the key technical challenges include: 1) the deemphasized role of semantic content that may be more important than low-level features in determining media quality; 2) the difficulty to optimally utilize cross-feature information for media quality analysis; and 3) the instability of the biologically/psychologically inspired features in reflecting human perception, and the lacking of a benchmark platform to evaluate the performance of these features.

This special session will target the most recent technical progresses on computational models for image, video and audio quality prediction, such as photo/video aesthetic quality ranking and photo cropping/retargeting. We also aim at discovering new types of visual/acoustic cues in computational quality models. The primary objective of this special issue is to foster focused attention on the latest research progress in this interesting area. We solicit original contributions which address the challenges from the computational models for visual /acoustic quality prediction. This special issue targets the researchers and practitioners from both the industry and academia. The topics of interest include, but are not limited to:

- New computational models for media quality evaluation, such as videos and music;
- Aesthetic models for various media enhancement techniques;
- Video and image summarization based on computational quality models;
- Different semantic models for multimedia quality prediction;
- Discovering low-/high-level visual features for multimedia quality prediction;
- Visual aesthetics prediction for multimodal applications;
- New feature fusion/selection techniques for multimedia analysis;

- Multimodal quality models for event and abnormal detection;
- Visual quality prediction for photo and video management systems;
- Human interactive learning for multimedia quality prediction;
- Video/audio quality prediction by mimicking human perception;
- Computational quality models for large-scale multimedia retrieval;
- Datasets, benchmarks and validation of visual quality of experience;
- Discovering advanced descriptors for evaluating multimedia quality.