# Introduction to the Special Section on Objective Quality Assessment of Speech and Audio

MODERN telecommunication networks are becoming more and more complex, and carry, besides the traditional narrowband speech signals (about 3.5-kHz bandwidth), wideband speech signals and other audio signals. In addition to existing traditional public switched telephone networks, various types of mobile and internet-based networks are widely used. The resulting interconnected heterogeneous network increases, significantly, the number of factors that can affect speech and audio quality; hence, understanding the impact of individual system components and combinations of them on the end-to-end speech and audio quality is difficult. In addition, these newer networks provide a tradeoff between service quality and cost. Thus, the reliable estimation of speech and audio quality over modern telecommunication networks is very important, not only for network systems design and development, but also for the maintenance of quality of service (QoS).

Historically, formal subjective listening tests have been used in evaluating the quality of speech and audio processing and transmission systems such as speech and audio coders. These tests require well-controlled facilities and expertise to provide reliable results related only to quality. Throughout the development and deployment of network systems, it is necessary to investigate the impact of specific system components, combinations of them, and sets of system parameter values on the perceived quality of speech and audio. Since it is difficult to obtain these results promptly and repeatedly by subjective tests, which are time consuming and costly, it is desirable to have a computational model that can reflect subjective quality ratings in reliable manner.

Over the last three decades, many objective speech and audio quality estimation models have been proposed. The goal of objective models is to mimic human performance in quality ratings. As a result, the research effort has been on understanding how the human auditory system processes speech and audio signals. This is a challenge as we do not yet have enough knowledge to establish a human-like objective model. Despite this difficulty, there has been significant progress in the past few years. However, the real-world experience obtained with the more successful and standardized measures has also indicated shortcomings and remaining challenges.

Digital Object Identifier 10.1109/TASL.2006.885521

The origin of this special issue lies in a desire to provide better insights into the recent advances and challenges in objective quality assessment of speech and audio. The 35 submissions received in response to the call for papers confirmed this topic to be an active area of research as well as the timeliness of this special issue. After a thorough review process, 13 papers were accepted for this issue. The first paper by the guest editors introduces key concepts, challenges, standard techniques, and recent advances in the field as a whole. The rest of the papers address specific and detailed challenges: objective quality assessment models for speech and audio waveforms motivated by human auditory perception; modeling in quality estimation without using a reference speech signal for comparison; the extension of the ITU-T G.107 E-model to wideband speech and arbitrary packet loss conditions; parametric models considering the impact of edge terminals in voice over IP (VoIP); conversational quality assessment in VoIP; spatial audio quality for multichannel audio; and the application to monaural speech separation and performance evaluation of speech recognition systems.

We would like to thank the former and current Editors-in-Chief, Isabel Trancoso and Mari Ostendorf, for their help and support throughout this process. We thank all of the authors and reviewers for their efforts; we also thank Kathy Jackson from the Signal Processing Society for her assistance.

DOH-SUK KIM, *Guest Editor*
Lucent Technologies
Whippany, NJ 07981 USA

JOHN G. BEERENDS, *Guest Editor*
TNO Information and Communication Technology
2600 GB Delft, The Netherlands

ODED GHITZA, *Guest Editor*
Sensimetric Corporation
Somerville, MA 02144 USA

PETER KROON, *Guest Editor*
Agere Systems
Allentown, PA 18109 USA

ANTONY RIX, *Guest Editor*
The Technology Partnership (TTP)
Cambridge SG8 6EE, U.K.