

Question 6/12 – Analysis methods for speech and audio using complex measurement signals

(Continuation of Question 6/12)

1 Motivation

Terminal and network equipment increasingly includes complex signal processing techniques; super-wideband and fullband systems have entered the market place. Most devices cannot be regarded as linear, time-invariant systems. The subjectively relevant transmission characteristics of such equipment need to be correctly determined using adequate measurement methods. There is a need of having reproducible, well-defined measurement methods available for certification labs as well as for developers which ideally should be combined to one quality value.

Test signals and analysis techniques for use in telephony have been collected in previous study periods. This work led to updated Recommendations ITU-T P.340, P.501, P.502 and P.505. New test signals allow evaluating many different parameters more realistically and are no longer limited to narrowband and wideband. However there is still lack of analysis methods for mixed content such as speech and music. Modern speech codecs allow the transmission of signals of any kind. Existing methods and to some extent signals need to be adapted since they may no longer be appropriate for new signal processing methods. In addition the interaction of signal processing at various locations of a connection needs to be investigated more in detail.

The evaluation methodologies for speech and audio processing are still incomplete and need further improvement, new technologies in hands-free, conference systems, in-car communication and speech processing require the adaptation of existing testing methodologies and the study of new procedures. There is a need to produce new product-oriented Recommendations including hands-free functions as, mobile, IP, conferencing and audiovisual terminals.

The following major Recommendations, in force at the time of approval of this Question, fall under its responsibility:

P.50, P.59, P.300, P.310, P.311, P.313, P.330, P.340, P.341, P.342, P.381, P.382, P.501, P.502, P.505

2 Question

The following items are to be considered within the study of the Question, special consideration should be given super-wideband/fullband systems, to mobile terminal signal processing, to VoIP terminals and signal processing used in VoIP including maintenance of existing Recommendations:

- What kind of new complex signal processing used in terminals, systems and networks may influence speech and audio transmission quality and what objective testing methodology can be used?
- What kind of techniques can be used to simulate time-variant use and time-variant behaviour of telecommunication equipment?
- What additional type of test signals and testing techniques are needed for wideband, super-wideband and fullband transmission systems?
- What type of test signals and analysis procedures can be used for spatial audio?
- What test signals other than speech and noise are needed and how can they be defined?
- What test signals can be used for the simulation of noisy environments?
- What methods are suitable for the objective assessment of background noise transmission and to what extent can the background noise transmission be assessed without making reference to the background noise signal?
- What testing methods/signals can be used to optimize background noise transmission in combination with VAD and comfort noise insertion techniques?

- What testing methods/signals can be used for real-time signal processing techniques such as in-car communication (ICC)?
- What testing methods are needed for speech and audio enhancement devices and what are the limits for the different quality determining parameters identified?
- What are the consequences on the speech quality of speech processing implemented in hands-free terminals and new types of conferencing devices, e.g., Smart Home? What characteristics and limits can apply?
- What characteristics and limits can apply other speech processing techniques such as speech recognition systems?
- What are the implications of the interaction between terminal signal processing and network signal processing on speech quality?
- How can existing and/or new speech quality parameters be combined to a single speech quality representation covering all conversational aspects?

3 Tasks

Tasks include, but are not limited to:

- improve/adapt existing test signals and objective speech quality testing methodologies;
- identify and study new basic objective testing methodologies in telecommunications;
- identify and study new basic objective testing methodologies for audio;
- identify and study new basic objective testing methodologies for spatial audio;
- identify and study new testing methodologies for real time signal processing techniques used e.g., in ICC (in-car communication);
- identify and study new testing methodologies for background noise transmission quality;
- identify and study the impact of time-variant user behaviour and time-variant signal processing by defining new test methods and setups;
- improve testing methods for speech enhancement devices;
- add new testing methodologies, improve the existing testing techniques for modern hands-free and conference terminals;
- study applications to multichannel sound pick up (arrays) and multichannel/multi-device sound reproduction (incl. spatialization, stereo).
- maintenance of Recommendations P.300, P.310, P.311, P.313, P.341, P.342, P.381 and P.382.

An up-to-date status of work under this Question is contained in the SG12 work programme https://www.itu.int/ITU-T/workprog/wp_search.aspx?sp=17&q=6/12.

4 Relationships

WSIS Action Lines

- C2

Sustainable Development Goals

- 9

Recommendations

- P.79, G.161, G.168, G.169, P.1100, P.1110, P.1130, P.1140, P.370, P.380

Questions

- 4/12, 5/12, 9/12, 10/12

Study Groups

- ITU-T SG16

Other bodies

- ETSI TC STQ, 3GPP SA4, TIA, IEEE, IEC