Question 9/12 – Perceptual-based objective methods and corresponding evaluation guidelines for voice and audio quality measurements in telecommunication services

(Continuation of Question 9/12)

1   Motivation

The work of this Question will focus on objective, perceptual and mainly signal-based methods for evaluating quality parameters in telecommunication scenarios. Primarily, the methods under study should concentrate on user-perceived quality characteristics. Consequently, these methods and algorithms include perceptual approaches. They model results and procedures, which are applicable in subjective tests. So that subjective procedures will get an objective counterpart by using the same scaling and basic procedures.

An example for that is the successful standardization of Recommendations P.862, P.862.1, P.862.2, P.862.3 and the P.863 up to fullband audio, perceptual based methods those models objectively Listening Only Tests with Absolute Category Rating for the evaluating of the Listening Speech Quality according to Recommendation P.800. A no-reference counterpart of P.862 was approved as P.563.

This Question will extend the objective evaluation of Listening Quality – the main issue up to now – to other quality aspects of voice telephony like talking quality and quality dimensions in no-reference and full-reference setups, including perceptual, signal-based models for objective rating of multi-channel and spatial audio in telecommunication services. Under consideration of new generation telecommunication services, also other media than speech like music should be taken into account.

Furthermore, the evaluation of transmitted noise – especially after processing by noise suppression systems – should be covered by the work of this Question, the same as objective prediction of speech intelligibility. This Question analyses and recommends also methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models and gives guidance for developing quality prediction models in general and especially by means of machine learning and artificial intelligence.

This Question will also continue and finalize the ongoing work on P.ONRA, P.AMD/P.SAMD and P.MLGuide.

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

P.563, P.862, P.862.1, P.862.2, P.862.3, P.863, P.863.1, P.1401

2   Question

Study items to be considered include, but are not limited to:

–   An already defined work item is the objective assessment of talking quality. Therefore at first a reliable subjective test method has to be established. In a second step, an objective model can be developed.

–   In addition to the existing objective models like P.863 or P.563 that are producing single numbers describing the overall quality; a need for additional information about possible quality degradations and quality dimensions are requested by the market. This is studied under P.AMD (full-reference) and P.SAMD (no-reference).

- Furthermore, the objective assessment of audio signals such as music transmitted over telecommunication links like WCDMA, LTE and 5G with modern codecs and terminals should be investigated.

- The objective rating of the annoyance of noise and residual noise – especially by processing by VQE's – in voice communications has to be investigated. Here a close relationship to the subjective method P.835 is given. A study item P.ONRA is already launched in this Question.

- Perceptual, signal-based models for objective rating of multi-channel and spatial audio in telecommunication services are interesting under the scope of this Question.

- The determination of the quality of synthesized speech in an instrumental way, e.g. using the objective perceptual methods, is an interesting topic in this Question as well as methods for objective prediction of speech intelligibility.

- This Question analyses and recommends methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models. These statistics can be applied to objective prediction models which can be translated to an estimated subjective judgment of a dedicated subjective test procedure. This Question discusses frameworks, metrics and example procedures for those statistical analyses and reporting. Furthermore, this question gives guidance to develop quality prediction models in general and specifically by means of machine learning and artificial intelligence as in P.MLGuide.

3   Tasks

Tasks include, but are not limited to:

- maintenance and enhancement of P-series Recommendations with regards to objective quality testing methods and perceptual models as P.863, P.863.1 and P.563;

- completion of Recommendations on

  • objective estimation of individual quality dimensions as full-reference approach P.AMD and its no-reference counterpart P.SAMD;

  • objective evaluation of noise reduction systems (P.ONRA);

  • guidance for using machine learning techniques in prediction model development.

- development of a Recommendation for objective, perceptual quality prediction of non-speech signals (e.g. music) in telecommunication services;

- development of a Recommendation for perceptual, signal-based models for objective qualitative rating the perception of multi-channel and spatial audio in telecommunication services.

An up-to-date status of work under this Question is contained in the SG12 work programme
http://www.itu.int/ITU-T/workprog/wp_search.aspx?q=9/12

4   Relationships

WSIS Action Lines

- C2

Sustainable Development Goals

- 9

Recommendations

– P-series, G.100- and G.1000-series

Questions

– 4/12, 6/12, 7/12, 11/12, 14/12, 15/12, 16/12, 19/12

Study Groups

– ITU-T SG16

Other bodies

– ETSI TC STQ, 3GPP