



3SQM™

ADVANCED NON-INTRUSIVE VOICE QUALITY TESTING

OPTICOM GmbH
Naegelsbachstr. 38
91052 Erlangen
GERMANY

Phone: +49 9131 / 530 20 0
Fax: +49 9131 / 530 20 20
Email: info@opticom.de
Website: www.opticom.de

Further information:
www.psqm.org
www.pesq.org
www.3sqm.org

White Paper by OPTICOM GmbH, Germany

CONTENTS

1	EXECUTIVE SUMMARY	3
2	INTERNATIONAL STANDARDS.....	4
2.1	P.SEAM - Non-Intrusive Voice Quality Analysis	4
2.2	P.AAM - Acoustic Extension of PESQ.....	4
3	INTRUSIVE VS. NON-INTRUSIVE TESTING.....	6
4	PERCEPTUAL TESTING VS. QUALITY ESTIMATION MODELS.....	8
4.1	The E-Model.....	8
4.2	Quality Estimation based on VoIP Protocol Information	8
4.3	Perceptual Modeling of Listening Quality.....	9
5	THE 3SQM™ VOICE QUALITY ANALYSIS.....	12
5.1	General Overview	12
5.2	Structure of the 3SQM™ algorithm	13
5.3	Preprocessing	13
5.4	Basic Distortion Classes and Speech Parameter Extraction	14
5.5	Detection of Dominant Distortion	14
5.6	Final Quality Estimate	15
6	PERFORMANCE RESULTS	16
7	PRODUCT AVAILABILITY.....	18
7.1	Stand-alone OPERA Products.....	18
7.2	OEM libraries	18
7.3	Integrated Network Management Systems.....	18
8	ABOUT OPTICOM.....	20
9	REFERENCES.....	21

1 EXECUTIVE SUMMARY

Since 2001, PESQ (ITU-T P.862, [2], [3], [26]) forms the state-of-the-art technique and international standard for advanced perceptual voice quality analysis. PESQ is an intrusive voice quality test, applicable to assess the end-to-end quality of next generation networks, based on simulating a subjective listening test.

OPTICOM has a long track record in the design, marketing and licensing of perceptual audio quality test algorithms and products. The new Single Sided Speech Quality Measure 3SQM™ represents the joint development of a new ITU-T standard (ITU-T P.563, [27]) for advanced non-intrusive voice quality testing. 3SQM™ allows for accurate voice stream analysis using perceptual criteria while being able to be applied to any real-world voice conversation. The technology behind 3SQM™ which was developed in a leading consortium together with Psytechnics and Swissqual is – like PESQ, based on a generic perceptual approach and therefore independent from the network technology being assessed. The underlying technology was released by the ITU-T in May, 2004 as new ITU-T recommendation P.563. This new standard does not supersede intrusive analysis, such as PESQ, but it marks the future industry standard in non-intrusive voice quality testing.

2 INTERNATIONAL STANDARDS

Currently, international standardization is ongoing within the [International Telecommunications Union \(ITU\)](#). A major extension is developed within Question 9 of Study Group 12, covering the **acoustic extensions** to ITU-T P.862/PESQ, the state-of-the-art ITU-T standard for intrusive voice quality testing.

Based on the unprecedented success of a joint development, such as **PEAQ** [4][5][13], the ITU-R standard for Perceptual Evaluation of Audio Quality, OPTICOM could again successfully assemble two leading industry consortia to join expertise in the development of the two new ITU-T recommendations. While the acoustic extensions are still under development the new non-intrusive voice quality measurement with the working title P.SEAM became the standard P.563 within the ITU-T.

2.1 P.SEAM - Non-Intrusive Voice Quality Analysis

Under the working title "**Single Ended Assessment Model**" (**P.SEAM or P.563**), the ITU-T had channeled standardization of various proprietary proposals to estimate voice quality non-intrusively meaning that the measurement takes place at the listener's side, only. No reference signal has to be inserted into the network for this purpose. At the trade-off of losing some accuracy compared to an intrusive measurement technique, such as P.862/PESQ, the new single ended measure provides the terrific advantage that it is able to measure at almost any point in the network with any real-world speech signal. Also, the single-ended measurement according to P.563 is not restricted to certain reference signals, which means, that it can be applied to any real-world telephone conversation. P.563 was approved as an international standard by the ITU-T committee for non-intrusive voice quality measurements in May, 2004 [27].

Being one of the three proponents, OPTICOM has proposed key technology derived from P3SQM which was originally developed by KPN Research (now TNO). The powerful consortium includes further Psytechnics and Swissqual, thus representing the know-how of the leading experts and co-developers of several perceptual audio quality test algorithms, including PSQM, PAMS, PESQ and PEAQ. The consortium joins an impressive number of pending and approved patents on the technology of voice quality testing fundamentally evidencing their huge expertise on that topic.

OPTICOM's implementation of the novel non-intrusive measurement standard P.563 is introduced under the brand name 3SQM™.

2.2 P.AAM - Acoustic Extension of PESQ

Under the working title "**Acoustic Assessment Model**" (**P.AAM**), this new development will provide an extension for the current [ITU-T recommendation P.862/PESQ](#) including now also acoustic interfaces. PESQ provides for end-to-end quality testing of voice-band signals at electrical interfacing to the network components. The newly devised extension will also support enhanced acoustic testing functionality for terminals to include hand-sets, head-set and hands-free kits in the measurement. It is expected that the

acoustic version which is highly based on the original P.862/PESQ model, will most likely become the complementing new standard P.863. It will be the first choice for advanced test labs, who do not spare the test efforts needed for acoustic setups, while P.862/PESQ will continue to form the state-of-the-art base line voice quality measurement for most users.

In this impressive consortium, OPTICOM was responsible for the code integration of the other partners, who are Deutsche Telekom, T-Nova (Berkom), KPN Research (now TNO), and Psytechnics.

Being one of the developers and a party to each of the consortia, OPTICOM will be in the position to be one of the first to release products and OEM technology based on the upcoming new standard.

3 INTRUSIVE VS. NON-INTRUSIVE TESTING

Intrusive test methods, like PSQM [25] and PESQ [26], insert a reference signal into the device under test. Like in a subjective test, the evaluation is based on a **natural voice or music sample**, typically of few seconds duration. A stored reference is sent through the device under test, and the received listening quality is analyzed by comparing the recorded sample to the original. Using natural voice or music signals for the measurement is superior to applying artificial test signals, such as sinusoidal tones or noise, as the latter ones do not properly model the signal characteristics of a normal operation.

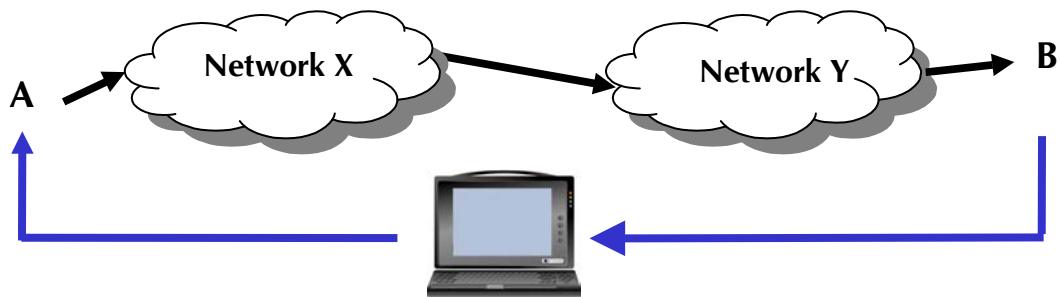


Figure 3.1 A typical setup for an **intrusive test**: The test system sends a reference speech stimulus that is inserted into a network connection at point A (origin), while the received signal at point B (termination) is fed back to the test system for difference analysis.

However, due to the fact that the reference signal has to be inserted into the device under test, such measurements are often referred to as **'intrusive'** measurements. That is, for a telecom application a test system like OPTICOM's OPERA will generate test calls. This could lead to complex setups in the case of widely distributed networks. Multiple network test setups are needed at various locations, and they talk to each other through an IP connection in order to control synchronized measurements (see figure 3.2).

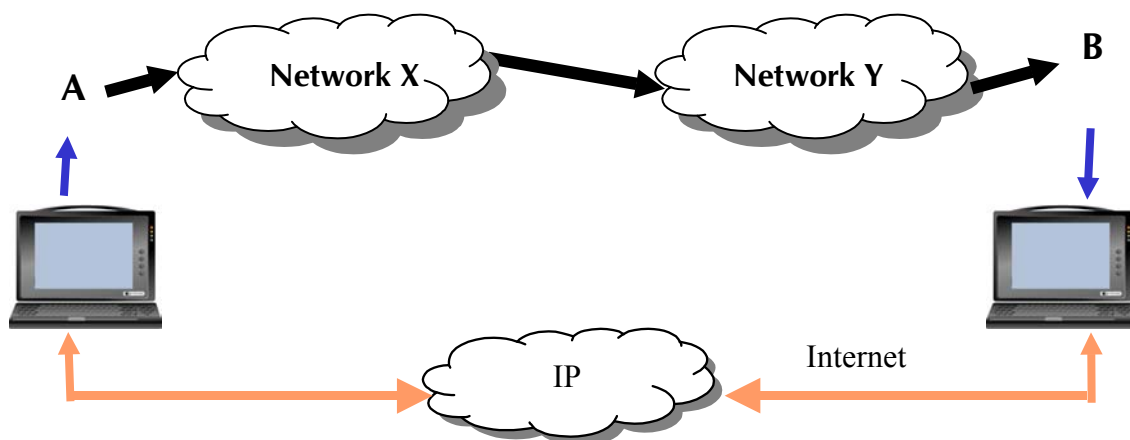


Figure 3.2 A real-world implementation of an **intrusive test**, e.g. based on OPTICOM's OPERA testers: At point A (origin) a test system will setup the call and will insert a reference signal into the network, while at point B (termination) another test system will acquire the signal under test and will perform the PESQ analysis.

From the perspective of a network operator who is interested in a permanent network control a '*non-intrusive*' method, only based on single sided monitoring without generating extra traffic may be preferable. Such measures are available, too, but due to the missing information of the source signal, they are not as reliable and accurate as intrusive measures. On the other hand they still can be employed to derive a reasonably accurate quality indicator. Most likely non-intrusive test methods will not supersede intrusive analysis, nonetheless a fertile co-existence of both measures is expected in the future.

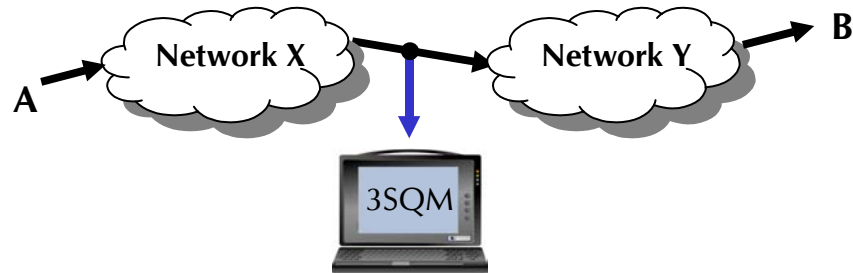


Figure 3.3 Non-intrusive test methods can be employed at any point in the network

Note:

A non-intrusive measurement, like **3SQM™** may continuously be applied for permanent network quality monitoring. In the fault case, which results most probably in a considerably decreased 3SQM™-MOS, an engineer can further analyze the cause of the problem by employing an intrusive measurement, like **PESQ**, to get more accurate and detailed results for advanced diagnostics and trouble shooting.

4 PERCEPTUAL TESTING VS. QUALITY ESTIMATION MODELS

4.1 The E-Model

The **ETSI E-Model** as defined in ITU-T G.107 [16] is a planning tool that assigns a certain equipment impairment factor I_e to each piece of equipment in the transmission chain. These I_e values are summed up and combined with several other parameters to form the final **R** factor or **R rating**. This **R rating** is a coarse estimate of the quality that can be expected if the network is realized in that way it was planned.

Although the E-Model is an excellent planning tool, it can never replace real measurements on the final network, since it has to make some very wide ranging assumptions. **R** ranges from 0 for perfect up to 100 for terrible voice quality. Note that there is a well defined relation between **R** and the MOS score. To allow for the comparison between the estimates from the network planning phase and the QoS of the live network, PESQ implementations, as in OPERA, provide the **R rating** as well. It is directly derived from the overall MOS as it is calculated by PESQ. It neither takes delay nor echo nor attenuation into account and consequently should be considered more closely corresponding to the G.107 I_e value than to the **R** factor. In fact **R** is introduced as a conversational measure, rather than a listening quality index [16].

Due to the fact that the E-Model is relying on many assumptions it can therefore only produce an estimate of the overall voice quality. In order to take this into account, the novel supplement ITU-T P.800.1 [20] defines a new language, which must be used in the context of the E-Model to pinpoint the provenance of the reported values: **MOS-LQE** (which stands for Listening Quality *Estimate*) versus **MOS-LQO** (meaning Listening Quality *Objective* Measure, e.g. with PESQ).

4.2 Quality Estimation based on VoIP Protocol Information

Although primarily developed as a pure planning tool, the E-model quality estimation approach has been implemented by some vendors to develop lightweight algorithms, e.g. for VoIP quality estimation. For instance, by carefully monitoring the jitter buffer behaviour, one can find out about packet loss and time varying effects, like varying delay. Based on these physical parameters, an R rating can be calculated. This information is of course limited as it can only characterise the performance of the individual network component. Consequently, to derive a quality estimate for the end-to-end listening quality within the network, one must not only transfer this piece of information through the network and gather similar information from all other network components, but one must also know about the non-linear interaction between these artifacts. As a presumption, it is therefore applicable to a homogenous network only and requires full access to such a network. Only under these circumstances a reasonable quality estimate can be expected. It is however obvious that in a heterogeneous network environment, for example as shown in figure 4.1 the assessed physical parameters, for instance of the VoIP part of the network are only of

limited influence on the total call quality. This is especially true if the network carries voice which was re-encoded several times with different speech coding schemes, for example in cascaded mobile, fixed and VoIP networks.

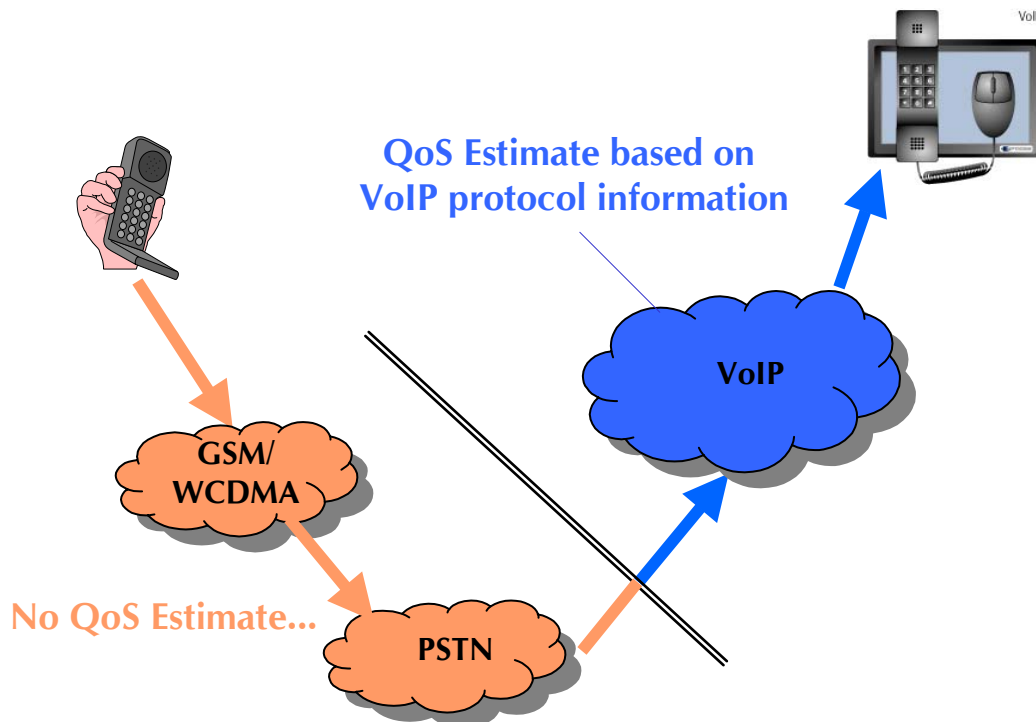


Figure 4.1 QoS estimates which are based on protocol information are limited to homogenous networks, e.g. VoIP and have no real knowledge of the voice signal quality

In the case of a non-intrusive voice quality analysis, such as 3SQM™, even heterogeneous networks can be accurately assessed by analysing the real voice stream, as perceived by the customer.

Nevertheless, further work is going on in the ITU-T under Question 16/12 ("In-service non-intrusive assessment of voice transmission performance") [18] with the scope of standardization of a lightweight protocol information only based quality estimate.

4.3 Perceptual Modeling of Listening Quality

The design of objective measurement methods based on human perception goes back to the eighties. It is based on the research work of Zwicker, Schröder, Brandenburg et al. The first algorithm that was implemented into a real measurement system was NMR (Noise to Mask Ratio) in 1989. The best known algorithms in the past were PAQM, PSQM[25], NMR, PERCEVAL, DIX, OASE, POM. Except for PSQM, all of these algorithms were developed to assess the quality of wideband audio codecs. This is due to the fact that the widespread use of perceptual codecs started earlier in the broadcast environment than it did in telecommunications. In 1996 PSQM was standardized as ITU-T Rec. P.861 for speech quality measurement. It showed superior correlation with subjective tests compared to all the other proposals that were not based on human perception. Contrary to PAQM, PSQM, NMR, PERCEVAL, DIX, OASE and POM, PEAQ [5][13] was developed as a joint collaboration. PEAQ was standardized in 1998 as ITU-R Rec. BS.1387 for wideband audio testing. With the ongoing development of speech coding, especially for packet transmission, new algorithms for speech quality measurement were developed, like PSQM+,

PSQM99, MNB, PAMS, TOSQA, PACE and VQI. Verification tests performed by the ITU showed that PSQM99 was far better than the other proponents' algorithms. The second best was PAMS, but none of these proposals was good enough for a revision of the P.861 standard. Consequently PESQ was developed and standardized in 2000 as ITU-T Draft Rec. P.862 [26].

When comparing all of the relevant measurement algorithms they can be broken down to a block diagram as shown in figure 2. Although they all share the same basic structure they differ significantly in the way they try to model human perception.

The basic structure consists of two inputs: One for the (unprocessed) reference signal and another for the signal under test. Latter input signal may for example be the output signal of a codec that is stimulated by the reference signal.

In a first signal processing step the peripheral ear is modelled ("perceptual model", or "ear model") [7][8]. Of course, the implementations of the peripheral ear model differ widely between the various algorithms. In general it can be said that for wideband audio signals this part of the algorithm is more important than for speech quality measures and therefore it must be modelled more accurately as in PEAQ. In addition it can also be observed that there are significant improvements between the initial algorithms like PAQM or NMR and the latest developments like PEAQ. PEAQ probably uses the most accurate and most detailed perceptual model that has ever been implemented until today.

In a consecutive step, the algorithm models the audible distortion present in the signal under test by comparing the outputs of the ear models. The outputs obtained by this process are called MOVs ("Model Output Variables") which are useful for a detailed analysis of the signal.

The final goal is deriving a quality measure consisting of a single number that indicates the audibility of the distortions present in the signal under test. To achieve this some further processing of the MOVs is required which models the cognitive part of the human auditory system. Again various proposals exist for this step. They range from algorithmic descriptions (e.g. PESQ) to artificial neural networks (e.g. PEAQ). Most algorithms require time aligned input signals, however the process how to achieve this is usually not part of the model description. Just now with the new speech quality measures like PESQ, the delay compensation is an integral part of the model.

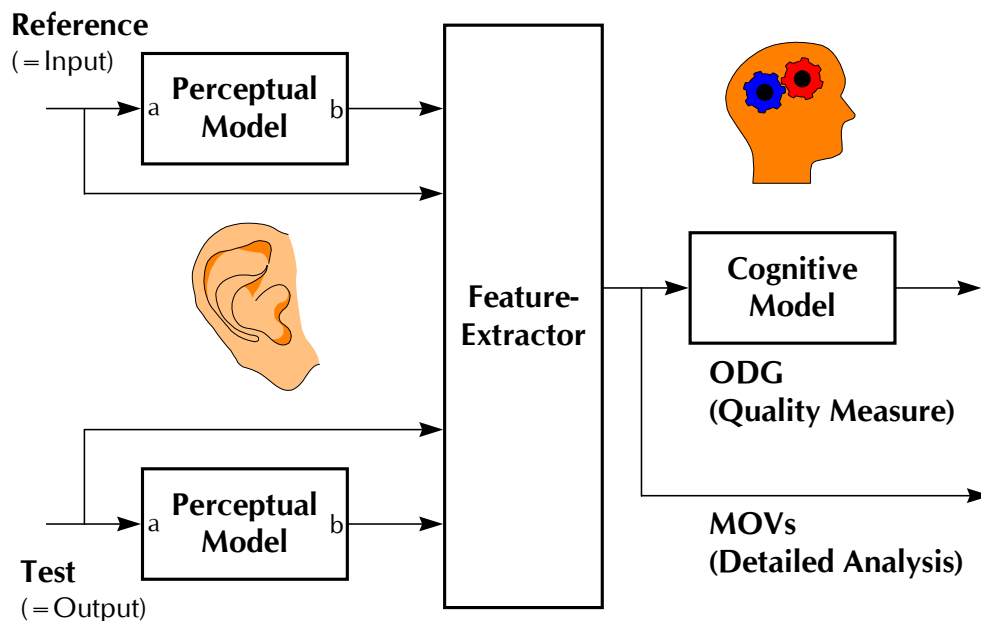


Figure 2: The structure of the generic perceptual measurement algorithm

Summary:

We can note that objective testing of voice quality based on perceptual techniques works, because it analyses **the transmitted voice signal** by modelling both, **the human ear** (perceptual modelling) and **the judgement behaviour** of a test subject (modelling the brain).

5 THE 3SQM™ VOICE QUALITY ANALYSIS

5.1 General Overview

Non-intrusive assessment of voice quality as known today can be based on two fundamentally different principles.

The first principle is looking at the signal processing to which the voice signal was exposed during the transmission, and makes assumptions on the amount of distortions introduced by the processing. The voice signal itself is not taken into account. Generally this type of algorithm can only be used with a priori knowledge of the exact transmission path and all equipment that is used in between the two endpoints of a communication path. As soon as heterogeneous networks are used, a call has to pass through foreign transit networks or the call routing is unknown, this type of assessment will fail. Frequently, also special equipment is required which traces the signal processing in routers, switches etc. Such measures are currently proposed for standardization for the assessment of pure VoIP networks. However the advantage of such metrics is that they are computationally slightly less expensive than other methods. Typical examples for such algorithms are VQMon and PsyVoIP.

The second approach is much more universal, since in contrary to the aforementioned metrics it analyzes the voice stream and not the transmission path. Here it is possible to assess any kind of voice signal without restrictions on the network or equipment type used. Such measures are applicable in any scenario, whether the call routing is known or unknown and independent from the signal processing used. Also, no modification of existing switches etc. is required if such a metric shall be deployed, since the only required information is the speech signal itself which is available **at any point in the network**. Also, such metrics do not make any assumptions on the amount of distortion introduced by the network. Moreover they measure the audibility of such distortions. Measures following this approach are typically built on very general models of the human vocal tract to model the speech generation, as well as psychoacoustic models to simulate the human hearing process. These measures are – though still very efficient - slightly more complex than those relying on protocol information only, but far more flexible in their applicability. In today's heterogeneous networks this is the only type of non-intrusive measurement that can be used with hardly any restrictions.

5.2 Structure of the 3SQM™ algorithm

3SQM™ is based on the second generic approach. It combines the essential parts of three independent and fundamentally different algorithms that were proposed earlier and which will be described in the next sections.

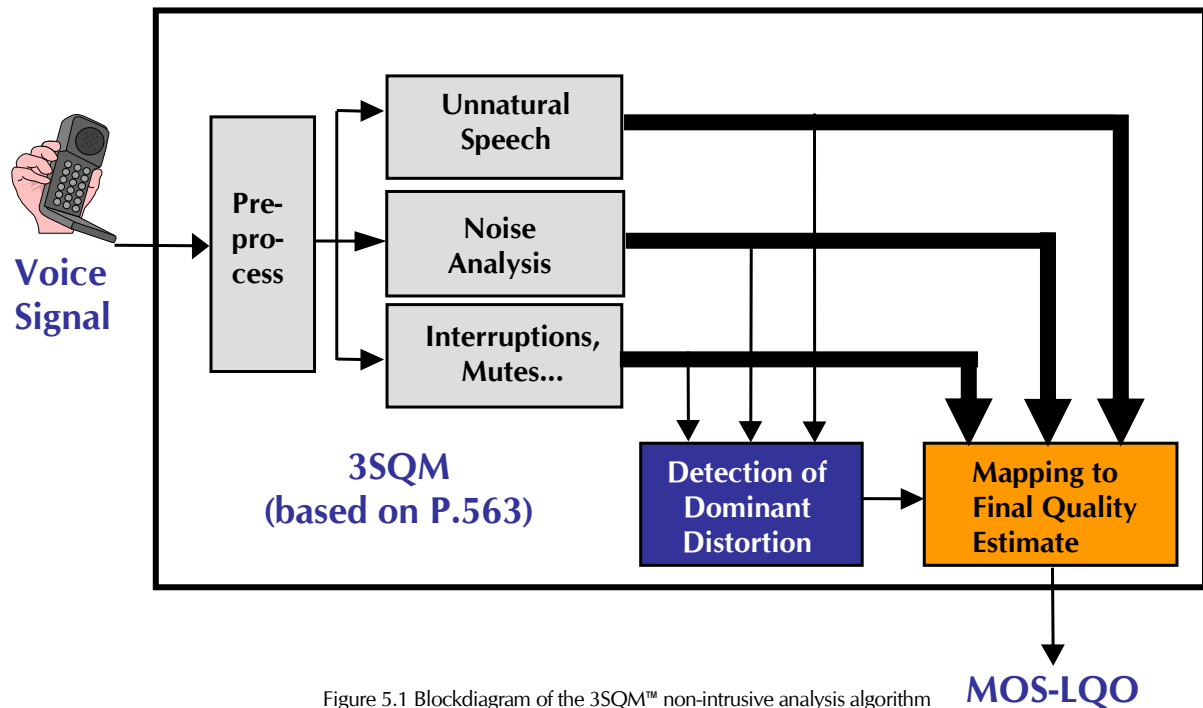


Figure 5.1 Blockdiagram of the 3SQM™ non-intrusive analysis algorithm

5.3 Preprocessing

Before the voice signal can be assessed properly it needs to be preprocessed in a first step. The important steps of preprocessing are:

- IRS receive filtering: The employed filter simulates a standard handset used in the laboratories for the subjective listening tests.
- Speech level adjustment.
- Separation in voice and non-voice parts via Voice Activity Detection (VAD)

5.4 Basic Distortion Classes and Speech Parameter Extraction

In a second stage the distortion and speech parameters are extracted for the speech signal. They are divided up into three main functional blocks which also correspond to the in recommendation P.563 considered main distortion classes. The main distortion classes are defined as:

1. ***Vocal tract analysis and unnaturalness of speech***
 - Basic speech quality depending on whether the talker is male or female
 - Robotic voice, e.g. caused by band limitation in GSM networks and unnatural voice like beeps
2. ***Analysis of strong additional noise***
 - Low static SNR (Background noise floor)
 - Low segmental SNR (Noise that is related to the signal's envelope)
3. ***Interruptions, mutes and time clipping***
 - Impairments as a result of lost packets in packet based transmission systems

All of these classes are based on very general principles which make no assumptions on the underlying network or distortion types occurring under certain conditions. The only prerequisite is the scientific knowledge on how human speech is generated and how it is perceived by human beings. This knowledge is built into the distortion model and does not vary with the application.

5.5 Detection of Dominant Distortion

During the workings for the standardization of P.563 the developers found, that several output parameters can be clustered to define single isolated distortion classes (see previous subsection). This models the phenomenon that any human listener focuses on the foreground of the signal stream. That is the listener would not judge the quality of the transmitted voice by a simple sum of all occurred distortions but because of a single dominant noise artifact in the signal.

Those distortion classes can be identified from a subset of the extracted parameters (see **Figure 5.1**) and are then prioritized according to the distortion's relevance with respect to the average listeners' opinions. The dominant distortion classes used with 3SQM™ are:

- **Low static SNR:** Occurs with a high background noise level.
- **Mutes:** Loss of packets in packet based transmission systems.
- **Low segmental SNR**
- **Unnatural voice – Robotization:** Highly periodic signal due to band limitation e.g. in GSM networks.
- **Basic speech quality:** In case, if the other models do not apply. Here two different models are used depending on whether the talker is male or female.

This part of the algorithm models the cognitive feature of human perception.

5.6 Final Quality Estimate

For each dominant distortion the model calculates the final quality estimate based on a selection of the MOVs. This quality estimate is equivalent to a MOS-LQO (Objective Listening Quality) value (1 is bad, 5 is excellent) according to P.800.1 and has a very high correlation with subjective listening test results. High correlations between objective and subjective tests are necessary as they prove the generally good reliability of the objective measurements, that is the model predicts the listeners judgment well. The correlations can be further improved with help of a non-linear mapping function. Often a third order polynomial function is employed that handles the non-linear edges of the MOS-scale. The non-linear property of the mapping function is necessary as it reflects the fact that verbal characterization ('excellent', 'good', ..., 'very annoying') translated to a numerical scale (5, 4, ..., 1) is not linear either.

6 PERFORMANCE RESULTS

In the following diagram the performance of the new, non-intrusive 3SQM™ analysis is compared to an intrusive analysis based on ITU-T P.862/PESQ. Please note that the correlations between objective and subjective results are shown per database for both analysis methods. It is amazing to see that for the number of 18 ITU subjective databases, the 3SQM™ performance is always above a correlation of 0,80 and in many cases it comes very close to PESQ's accuracy. Keeping in mind the much higher versatility of the non-intrusive approach, the newly approved ITU-T standard P.563 definitely marks a new milestone for perceptual voice quality testing. Further details of the databases used for this evaluation are shown in table 6.1.

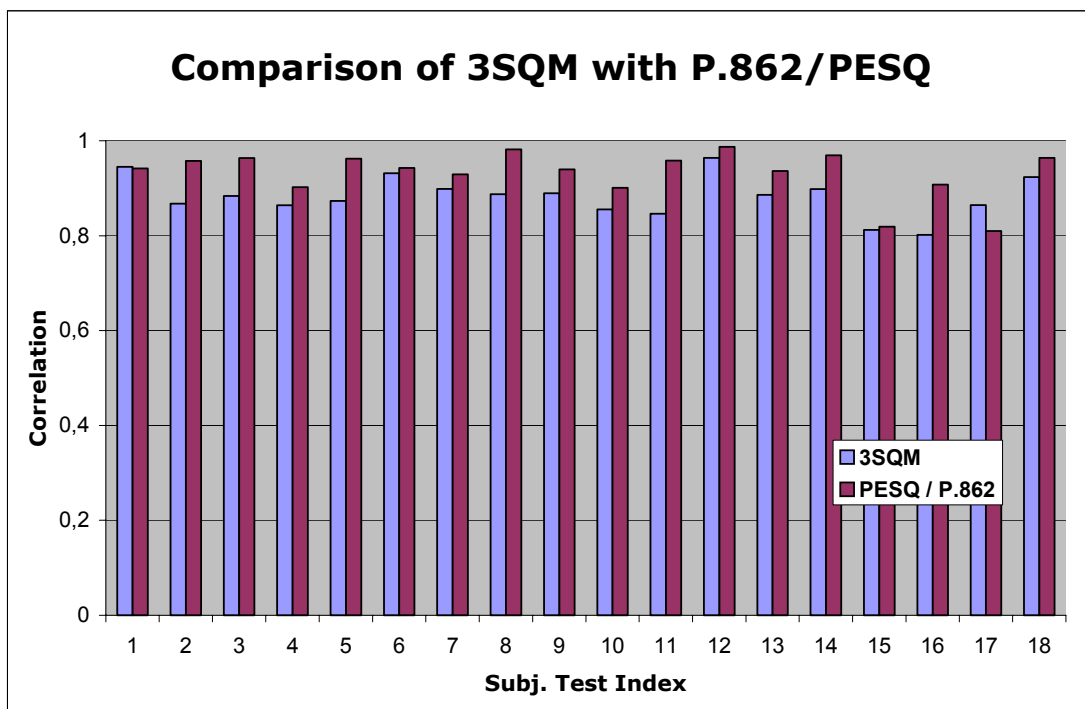


Figure 6.1 Correlation Results of 3SQM™ with real subjective tests, compared to results achieved with P.862/PESQ.

Subjective Test Databases
ITU Sup.23 expt.1: interworking with standards, CNET, French
ITU Sup. 23 expt.1: interworking with standards, NTT, Japanese
ITU Sup. 23 expt.1: interworking with standards, BNR, American English
ITU Sup. 23 expt.3: channel errors and noise, CNET, French
ITU Sup. 23 expt.3: channel errors and noise, CSELT, Italian
ITU Sup. 23 expt.3: channel errors and noise, NTT, Japanese
ITU Sup. 23 expt.3: channel errors and noise, BNR, American English
Q13 Ascom proponent test 1, Ascom, French
Q13 Ascom proponent test 2, Ascom, French
Q13 Berkomp proponent test, DT, German
Q13 Berkomp frame erasure test, DT, German
P86x ETSI VoIP measurement test, DT, German
Q13 BT Ylq test: codecs, errors, transcodings, noise, BT, British English
P86x Background Noise test English, BT, British English
P86x Network Emulation Dutch, KPN, Dutch
P86x Network Measurement Dutch, KPN, Dutch
P86x Network Emulation English, Ascom, British English
P.SEAM, GSM life Network, OPTICOM, German

Table 6.1 Real subjective test databases used for the comparison

7 PRODUCT AVAILABILITY

Being one of developers and a party in each of the consortia, OPTICOM is in the position to be one of the first to release products and OEM technology based on the new standards.

7.1 Stand-alone OPERA Products

It is expected that OPTICOM releases 3SQM™ in Q4/2004 as an additional software plug-in to both, the OPTICOM OPERA stand-alone testers as well as the OPERA Software Suite. OPERA 3SQM™ will add the non-intrusive capability to OPTICOM's general purpose signal quality analyzer that today marks already the reference for PEAQ and PESQ perceptual measurements.

7.2 OEM libraries

In addition to the stand-alone OPERA products, OPTICOM also added advanced 3SQM™ libraries for various common platforms to its portfolio of OEM libraries, available for licensing. An attractive licensing model will be available in the near future ensuring for a fast time-to-market for OPTICOM's OEM partners. It is expected that the licensing model will be composed of per unit or per channel fees, thus offering a flexible and largely scalable usage. Today an increasing number of 30+ well known industry players, including the Who is Who of the T&M manufacturers are counted to our OEM licensees.

7.3 Integrated Network Management Systems

In addition to the per unit based licensed use of OPTICOM's OEM libraries as above there will also be licensing terms available for enterprise wide usage or company internal licensing. This will be a compelling approach to add 3SQM™ to existing or newly deployed QoS management systems, even if you are not an equipment manufacturer. With its wide range of expertise, OPTICOM will also offer the services for integration and customer specific implementations.



Figure 7.1 The OPTICOM OPERA Voice/Audio Quality Analyser, offering PESQ, PSQM, PEAQ and soon also 3SQM™ analysis.

8 ABOUT OPTICOM

OPTICOM, the world leader in perceptual voice and audio quality testing solutions and the technologies provider of techniques such as PSQM, PSQM+, PEAQ, PESQ and 3SQM addresses the testing advantages of utilizing ITU's current and proposed standards for today's and future networks.

Under the mission statement "quality is our business ", OPTICOM focuses on top notch developments to gain for its customers improved quality in audio and video communications. With the new OPERA™ family of perceptual analyzers, the company proves it's worldwide reputation for state-of-the-art solutions to improve the audio quality of new media.

OPTICOM was founded by its President Michael Keyhl in 1995 as a "spin-off" company of the Fraunhofer-Institute, Germany's leading organization for applied research. OPTICOM's developers benefit from their broad experience in the research and development of perceptual based coding and evaluation techniques, such as MP3 and NMR, lasting back to the late 1980's.

Through many international contacts and cooperations with leading research organizations, OPTICOM has today gained an active role in the international standardization business, e.g. of the new ITU-R standard "PEAQ". OPTICOM is also continuously active in, or observing the work of the AES, EBU, ITU-T, ETSI, ISO/MPEG and others.

After being successful in business for more than four years, the company is growing fast and seeking to expand the number of their employees. OPTICOM is located in Erlangen, Northern-Bavaria, GERMANY, and has just recently opened offices and distributionship channels in the USA and Asia. For more information, please feel free to visit www.opticom.de.

9 REFERENCES

Literature

- [1] BEERENDS J. G., STEMERDINK J. A., *A perceptual speech quality measure based on a psychoacoustic sound representation*, J. Audio Eng. Soc., Vol. 42, No. 3, pp. 115-123, 1994
- [2] BEERENDS J. G., RIX A. W., HOLLIER M. P., HEKSTRA A. P., *Perceptual Evaluation of Speech Quality (PESQ) The New ITU Standard for End-to-End Speech Quality Assessment, Part I – Time-Delay Compensation*, J. Audio Eng. Soc., Vol. 50, No. 10, 2002
- [3] BEERENDS J. G., RIX A. W., HOLLIER M. P., HEKSTRA A. P., *Perceptual Evaluation of Speech Quality (PESQ) The New ITU Standard for End-to-End Speech Quality Assessment, Part II – Psychoacoustic Model*, J. Audio Eng. Soc., Vol. 50, No. 10, 2002
- [4] KEYHL M., SCHMIDMER Ch., WACHTER H., *A Combined Measurement Tool for the Objective, Perceptual Based Evaluation of Compressed Speech and Audio Signals*, 106th AES Convention, Munich, 1999
- [5] KEYHL M., SCHMIDMER Ch., Wachter H., Rath S., Stoll G., Colomes C., Sporer T., *Evaluating the Perceived Audio Quality (PEAQ) of Internet Audio Codecs*, 109th AES Convention, Los Angeles, 2000
- [6] MÖLLER S., BERGER J., *Describing Telephone Speech Codec Quality Degradations by Means of Impairment Factors*, J. Audio Eng. Soc., Vol. 50, No. 9 , 2002
- [7] ZWICKER E., FELDTKELLER R., *Das Ohr als Nachrichtenempfänger*, Hirzel-Verlag, Stuttgart, 1967
- [8] ZWICKER E., *Psychoakustik*, Springer-Verlag, Berlin - Heidelberg - New York, 1982

Standards

- [9] ETSI Technical Report ETR 250, *Transmission and Multiplexing (TM); Speech communication quality from mouth to ear for 3,1 kHz handset telephony across networks*, ETSI 1996
- [10] ISO/IEC/JTC1/SC29/WG11 Draft Document N1557, *Evaluation Methods and procedures for MPEG-4 tests*, 1997
- [11] ITU-R Recommendation BS.562-3, *Subjective assessment of sound quality*
- [12] ITU-R Recommendation BS.1116-1, *Methods for the Subjective Assessment of small Impairments in Audio Systems including Multichannel Sound Systems*, 1997
- [13] ITU-R Recommendation BS.1387-1, *Method for Objective Measurements of Perceived Audio Quality (PEAQ)*, Revised 11/01
- [14] ITU-R Recommendation BS.1534, *Method for the subjective assessment of intermediate quality level of coding systems*, June 2001
- [15] ITU-T Contribution COM12-74-E, *Review of Validation Tests for Objective Speech Quality Measures*, March 1996
- [16] ITU-T Recommendation G.107, *The E-model, a computational model for use in transmission planning*, May 2000
- [17] ITU-T Recommendation E.420, *Checking the Quality of the International Telephone Service – General Considerations*, 1988, (Extract from the *Blue Book*)
- [18] ITU-T Recommendation P.562, *Analysis and interpretation of INMD voice-service measurements*, May 2000
- [19] ITU-T Recommendation P.800, *Methods for subjective determination of transmission quality*, 1996
- [20] ITU-T Recommendation P.800.1, *Mean Opinion Score (MOS) Terminology*, March 2003
- [21] ITU-T Recommendation P.810, *Modulated Noise Reference Unit (MNRU)*, 1996
- [22] ITU-T Recommendation P.830, *Subjective Performance Assessment of Telephone-Band and Wideband Digital Codecs*, 1996
- [23] ITU-T Recommendation P.833, *Methodology for Derivation of Equipment Impairment Factors from Subjective Listening-Only Tests*, 2001
- [24] ITU-T Recommendation P.834, *Methodology for the Derivation of Equipment Impairment Factors from Instrumental Models*, 2002
- [25] ITU-T Recommendation P.861, *Objective Quality measurement of telephone-band (300 - 3400 Hz) speech codecs*, 1996

- [26] ITU-T Recommendation P.862, PESQ, *An objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*, February 2001
- [27] ITU-T Recommendation P.563, *Single-ended method for objective speech quality assessment in narrow-band telephony applications*, May 2004

**OPTICOM GmbH
Naegelsbachstr. 38
91052 Erlangen
GERMANY**

**Phone: +49 9131 / 530 20 0
Fax: +49 9131 / 530 20 20
EMail: info@opticom.de
Website: www.opticom.de**

**Further information:
www.psqm.org
www.pesq.org
www.3sqm.org**