INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.167

(03/93)

GENERAL CHARACTERISTICS OF INTERNATIONAL TELEPHONE CONNECTIONS AND INTERNATIONAL TELEPHONE CIRCUITS

ACOUSTIC ECHO CONTROLLERS

ITU-T Recommendation G.167

(Previously “CCITT Recommendation”)

FOREWORD

The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the International Telecommunication Union. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, established the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

ITU-T Recommendation G.167 was prepared by the ITU-T Study Group XV (1988-1993) and was approved by the WTSC (Helsinki, March 1-12, 1993).

____

NOTES

1 As a consequence of a reform process within the International Telecommunication Union (ITU), the CCITT ceased to exist as of 28 February 1993. In its place, the ITU Telecommunication Standardization Sector (ITU-T) was created as of 1 March 1993. Similarly, in this reform process, the CCIR and the IFRB have been replaced by the Radiocommunication Sector.

In order not to delay publication of this Recommendation, no change has been made in the text to references containing the acronyms "CCITT, CCIR or IFRB" or their associated entities such as Plenary Assembly, Secretariat, etc. Future editions of this Recommendation will contain the proper terminology related to the new ITU structure.

2 In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

CONTENTS

ACOUSTIC ECHO CONTROLLERS[1)]

(Helsinki, 1993)

1 General

1.1 Definition of acoustic echo controllers

acoustic echo controllers (denoted by AECs) are voice operated devices installed in audio terminals on the customer premises, used for the purpose of eliminating acoustic echoes and protecting the communication from howling due to acoustic feedback from loudspeaker to microphone.

1.2 Applicability of the Recommendation

This Recommendation is applicable to the design of AECs for audio terminals with digital or analogue line interfaces, and is intended for use in the following areas of telecommunications (denoted by applications):

- teleconferencing;
- loudspeaking (hands-free) telephones;
- videophone terminals;
- mobile and personal applications.

1.3 Objective of the Recommendation

The Recommendation specifies performance characteristics and values with which AECs must comply, and methods to verify these performances. The performances depend on the applications considered. Particular processing techniques are indicated as guidance for possible implementations, but are not compulsory.

1.4 Relevant Recommendations

The following Recommendations are relevant to the problem of acoustic echo control in audio terminals:

- Rec. P.30: Transmission performance of group audio terminals (GATs)
- Rec. P.31: Transmission characteristics of digital telephones
- Rec. P.34: Transmission characteristics of hands-free telephones
- Rec. G.131: Echo and delay
- Rec. G.173: Transmission planning aspects of the speech service in digital public land mobile networks.

In case of digital transmission, the standard digital speech coding formats considered are: G.711 (telephone band), G.722 (wide band). For mobile applications, the coding law is under study.

[1)] Values in brackets [ ] are provisional.

2 General definitions

2.1 Audio terminals

Audio terminals are designed for hands-free voice communication between individuals or groups of persons through analogue or digital networks. Any audio terminal is functionally accessed through user interfaces, network interfaces and test interfaces. The processing unit, electroacoustic transducers and attached circuits are internal parts of the terminal.

2.2 Interfaces

The interfaces are places (outside the terminal) or points (inside the terminal) where physical measurements can be done in order to adjust internal parameters of the terminal for proper operation and to verify performance. The items measured at the interfaces are used to infer the subjective quality that the local user(s) and the distant user(s) would experience.

2.2.1 User interfaces

There are two user interfaces:

– Receive interface (R_{out}): The place(s) where acoustic attributes relating to characteristics of speech listened to by the local user(s) are measured. This is also referred to as the position of the measuring microphone as described in Recommendation P.34.

– Send interface (S_{in}): The place(s) where acoustic attributes relating to characteristics of speech produced by the local user(s) are measured. This is referred to as the mouth reference point (MRP) in Recommendation P.34.

2.2.2 Network interfaces

There are two network interfaces:

– Receive interface (R_{in}): A point where the electrical signals received from the network are available.

– Send interface (S_{out}): A point where the electrical signals sent to the network are available.

If the terminal is connected to an analogue line, the interfaces shall comply with the characteristics specified in Recommendation Q.552. If the terminal is connected to a digital line (ISDN), the interfaces shall comply with the S-interface characteristics specified in Recommendations Q.554 and G.703[2)].

2.2.3 Test interfaces

The test interfaces are input and output points inside the terminal where signals and/or controls can be applied or measured for the purpose of performance verification.

2.3 Processing unit

The processing unit includes all the devices in the terminal performing signal processing functions on audio signals (except the devices which are parts of the electroacoustic transducers and attached circuits). A partial list of functions which may be in the processing unit is:

– A/D and D/A linear conversions of audio signals;

– signal processing for acoustic echo control;

– signal processing for other purposes (e.g. noise cancellation, room reverberation reduction);

– transcoding of audio signals between line code format and linear code.

---

2) In case of analogue line, the measuring equipment used for performance verification on the line side shall provide signal separation of both transmission directions by at least 60 dB for all frequencies in the transmission bandwidth (e.g. by using an adaptive hybrid in it). In case of digital line, this measuring equipment shall implement the same speech coding law as the one used in the terminal itself (e.g. Recommendations G.711, G.722, etc.) in the purpose to perform measurements on linear signals.

2.4 Electroacoustic transducers and attached circuits

The electroacoustic transducers are the loudspeaker(s) and microphone(s) attached to the terminal in normal operation. The attached circuits may include the amplifiers, switches, level adjustments and other devices which can be under the control of the user or automatically adjusted during the operation of the terminal like sound equalizers, etc.

The functional block diagram of a general audio terminal equipped with an acoustic echo controller is shown in Figure 1.

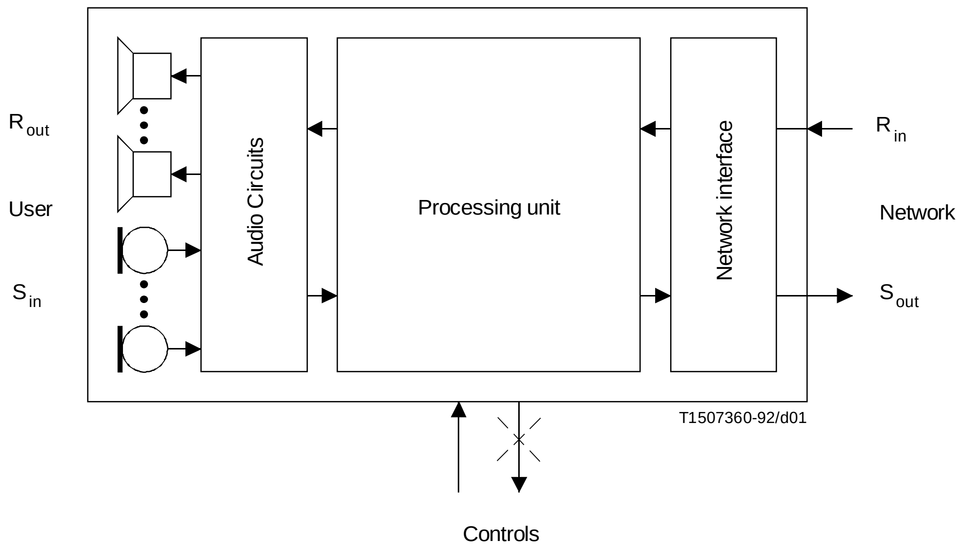

FIGURE 1/G.167

General audio terminal with AEC

3 Definitions relating to acoustic echo controllers

For the purpose of this Recommendation, the following definitions apply:

3.1 functional units: The functional units of an AEC are devices or parts of devices implemented in the processing unit, which contribute to the general function of acoustic echo control. There is no restriction on how to implement them. The following subclauses describe functional units which can be parts of an AEC (not a complete list).

3.1.1 acoustic echo canceller: A device which reduces the acoustic echo level with negligible effects on the local and distant users' speech. It is generally implemented by adaptive identification of the acoustic echo path response.

3.1.2 loss controller: A device which reduces the acoustic echo level by inserting variable losses on the received and/or transmitted audio signals.

3.1.3 non-linear processor: A device which reduces or cancels small echo signals by non-linear operation on the samples of the transmitted audio signal. A center-clipper is a typical device of this kind.

3.1.4 supplementary howling control device: A device which modifies some characteristics of the transmitted and/or received signals in order to improve the stability margin of the terminal. This function is typically implemented by an harmonic processor. To prevent network disturbances, such devices should be avoided in the case of terminals likely to be used on connections including network electric echo cancellers conforming to Recommendation G.165 which are not able to work properly with time-variant echo paths.

3.1.5 Summary

The functional units defined above can be combined for better performance. They can use all the available signals in the terminal (for example, the individual signals coming from several microphones arranged in an acoustic array). Moreover, they can be mixed with other functions (for example sub-band speech coding) for efficient implementation, provided that they do not modify the proper characteristics of these functions when they are operating.

It must be possible to enable, disable and reset functionally those devices when required by the test procedures described below in the Recommendation.

A functional block-diagram of a typical processing unit is shown in Figure 2.

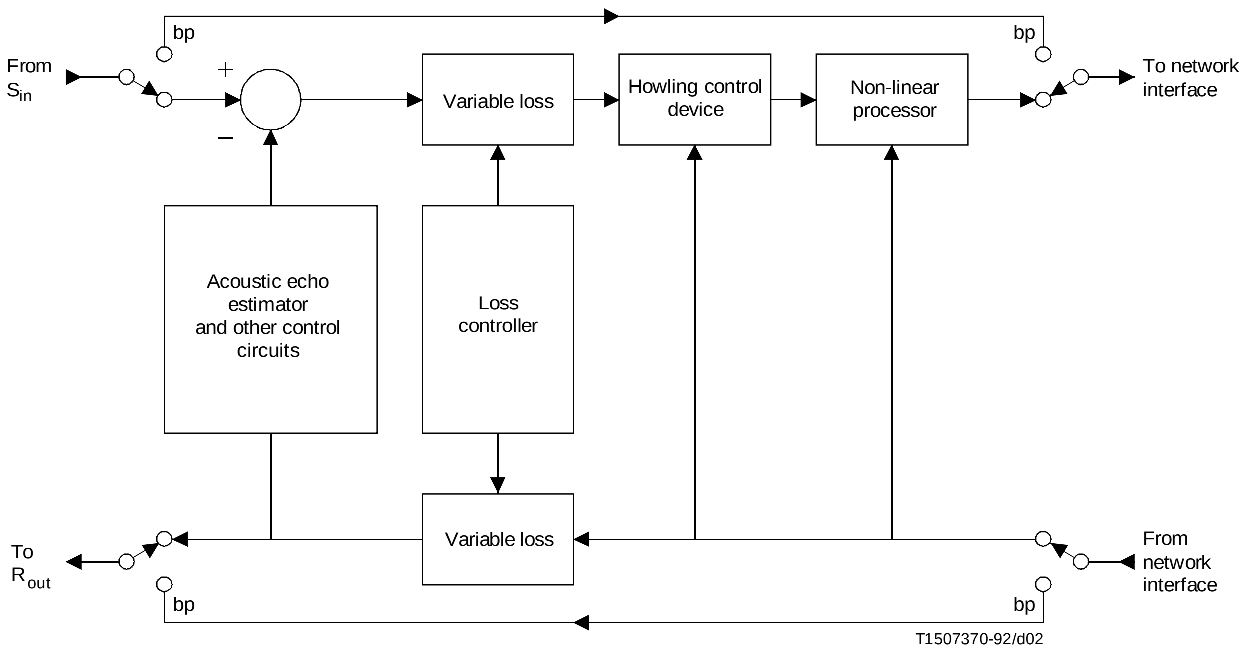

FIGURE 2/G.167

Functional block-diagram of a typical processing unit (AEC part)
(bp denotes bypass signal paths for testing purposes)

3.2 Items relevant to speech performance of acoustic echo controllers

The following subclauses are definitions of items which are specifically linked to the static and dynamic performance of AECs. Values and test procedures to measure these items are given in 5.

3.2.1 weighted terminal coupling loss – single talk (TCLwst): The weighted loss between the R_{in} and S_{out} network interfaces when the AEC is in normal operation, and when there is no signal coming from the local user[3].

3.2.2 weighted terminal coupling loss – double talk (TCLwdt): The weighted loss between the R_{in} and S_{out} network interfaces when the AEC is in normal operation, and where the local user and the far-end user are active simultaneously[3].

---

[3] The weighting is made according to the rule specified in Recommendation G.122 (computation of talker echo loudness rating). Care must be taken to avoid possible masking of singing effects by the weighting (under study).

3.2.3 received speech attenuation during double talk (Ardt): The received signal attenuation (at the R_{out} point) which is inserted by the AEC during double-talk events.

The frequency response on the receive side during double talk is under study.

3.2.4 sent speech attenuation during double talk (Asdt): The sent signal attenuation (at the S_{out} point) which is inserted by the AEC during double-talk events.

The frequency response on the send side during double talk is under study.

3.2.5 received speech distortion during double talk (Drdt): The total non-linear signal distortion at the R_{out} point which can be produced by the AEC during double-talk events.

3.2.6 sent speech distortion during double talk (Dsdt): The total non-linear signal distortion at the S_{out} point which can be produced by the AEC during double-talk events.

3.2.7 frequency shifting (or pitch ratio Pr): The shift-up of frequencies of the signal at the S_{out} and/or R_{out} points, due to howling control devices like harmonic processors (more details in 5.4.7).

3.2.8 break-in time – simple talk (Tonst): The time interval between the onset of the received signal (similarly the transmitted signal) and the instant when the attenuation on the receive path (similarly on the send path) reaches [3] dB. For this purpose, the other side is quiet.

3.2.9 break-in time – double talk (Tondt): The time interval between the onset of the received signal (similarly the sent signal) and the instant when the attenuation on the receive path (similarly on the send path) reaches the value Ardt (similarly Asdt). For this purpose, the signal in the opposite direction of transmission is held at a specified level.

3.2.10 initial convergence time (Tic): The time interval between the instant when a specified test signal is applied to the R_{in} port of the terminal (after all the functions of the AEC have been reset and then enabled), and the instant when the returned echo signal at the S_{out} port is attenuated by at least a pre-defined amount. The local user is not active.

3.2.11 recovery time after double talk (Trdt): The time elapsed between the end of a double talk event and the instant when the attenuation of the echo recovers a specified value (a signal is received continuously from the distant user).

3.2.12 terminal coupling loss during echo path variation (TCLwpv): The weighted attenuation of the echo which is observed during a specified echo path variation. The local user is not active[4)].

3.2.13 recovery time after echo path variation (Trpv): The time elapsed between the end of a specified echo path variation and the instant when the attenuation of the echo changes from TCLwpv to a specified (larger) value. The local user is not active.

4 Transmission specifications

4.1 Scope

These specifications are the basic requirements for AECs to transmit speech properly without noticeable quality degradation when included in audio terminals. AECs shall also meet specifications given in 5. Clause 4 covers all the applications mentioned in 1.2.

---

4) The weighting is made according to the rule specified in Recommendation G.122 (computation of talker echo loudness rating). Care must be taken to avoid possible masking of singing effects by the weighting (under study).

4.2 Application specifics

The AECs may be used in several different applications and the designers of terminals incorporating AECs should ensure that the network interface portion and the electroacoustic transducers and attached circuits of the terminal fully meet the appropriate Recommendations mentioned below. The measurements and requirements given in 4.3 to 4.10 are for the processing unit only.

4.2.1 Narrow-band speech

This applies to audio terminals connected to the PSTN and ISDN through telephone-band channels limited to frequencies under 3.4 kHz. For digital transmission the relevant Recommendation is G.712. For analogue transmission the relevant Recommendation is Q.552.

4.2.2 Wide-band speech

This applies to audio terminals connected to ISDN (e.g. high quality teleconference systems, wide-band hands-free telephones). The relevant Recommendations are G.722 and J.23 (except for the lower limit of bandwidth).

4.2.3 Other applications

For mobile radio and personal communication equipment the transmission characteristics specified in the corresponding Recommendations shall be used as references.

4.3 Measurements

The measurements shall be done separately on receive and send sides. On the receive side, a test signal is applied at the R_{in} point and the measurements are done acoustically at the R_{out} point. On send side, an acoustic test signal is applied at the S_{in} point and the measurements are done at the S_{out} point.

The following specifications refer to the difference between the measurements made when the processing part of the AEC is bypassed and when it is in normal operation.

4.4 Bandwidth

When activated, the processing in the AEC shall not modify the bandwidth of the terminal within the appropriate frequency mask. The bandwidth specifications are given in Recommendation G.712 for telephone band speech and in Recommendation G.722 for wide-band speech.

4.5 Attenuation distortion

The contribution of the AEC processing unit to the total attenuation distortion in the terminal shall be less than ± 1 dB within the appropriate bandwidth.

4.6 Delay

The values specified below correspond to the extra delay which can result from the AEC processing in the processing unit. The maximum delay permitted depends on the application. In any case, compliance with transmission planning objectives must be achieved. General information about transmission delays can be found in Recommendation G.114; Recommendation G.131 provides rules for echo control in the network.

For end-to-end digital communications (for example wide-band teleconference systems), the delay shall be no more than [16 ms] in each direction of speech transmission.

For hands-free telephones connected to the PSTN, the delay shall be no more than [2 ms] in each direction of speech transmission.

For videophones, the delay shall be roughly equivalent to the video codec delay (generally large) in order to achieve synchronization with lip movements. When using the videophone in PCM mode (i.e. without picture transmission), the delay shall be no more than [2 ms] in each direction of speech transmission.

For mobile radio systems, compliance with transmission planning objectives as specified in Recommendation G.173 shall be achieved. In any case, the delay shall be no more than [10 ms] in each direction of speech.

Other applications like cordless telephones, etc. are under study.

4.7 Delay distortion

The delay distortion added by the processing unit shall be no more than 1 ms for all the frequencies in the bandwidth of the system.

4.8 Non-linear distortion

The total non-linear distortion added by the processing unit shall be small in comparison with the distortion produced by the other parts of the terminal (under study).

4.9 Noise emitted by the AEC processing unit on the send side

The electric noise due to the AEC processing unit shall be such that the overall noise on the send side (including the out-of-band noise) will comply with the values specified in the corresponding Recommendations (P.30, etc). For this measurement, the microphones shall be switched off.

4.10 Acoustic noise produced by the AEC on the receive side

The contribution of the AEC processing unit to the acoustic noise coming from the loudspeaker(s) of the terminal shall be less than 1 dBA at the R_{out} point. The level of pure frequency components up to [16 kHz] should not produce a subjectively noticeable disturbance (associated value under study). For this measurement, the volume control of the terminal must be adjusted to its maximum value.

5 Acoustic echo control specifications

5.1 Scope

These specifications are the requirements for AECs to properly implement the acoustic echo control function while avoiding unacceptable degradations of speech coming from the distant user and from the local user. They cover all the applications mentioned in 1.2.

5.2 Measurement conditions

5.2.1 General

The measurements should be done as much as possible with the AEC operating in the terminal for which it has been designed. The transducers of the terminal should be installed at one or several typical positions in the room corresponding to its normal use. The level settings of the terminal should comply with the recommended values (see alignment procedures in Recommendations P.30, P.34, etc.). For example, in Recommendation P.30, the signal level at S_{out} is –22 dBV (± 2 dB) for an acoustic level at the MRP of –4.7 dBPa.

5.2.2 Measurement signals

5.2.2.1 Relevant Recommendations

Unless otherwise specified, the signals used for the tests are derived from Recommendation P.50 (artificial voice). The specification of measurement signals for AEC performance verification, especially for time-dependent characteristics like convergence, has been recognized as a difficult topic; one reason for this is that AECs can behave differently according to the kind of test signal used and of the particular sequence of this signal used for the measurement. The following choices of measurement signals are under study.

a) Use of a stationary noise with average spectrum complying to Recommendation P.50. Since this signal is stationary, the AEC performance should be less sensitive to the starting point in the sequence of signal than in the case of non-stationary signals like the artificial (non-stationary) speech signal P.50 and real speech. One must be aware that optimization of the AEC performance according to the spectral characteristics of this signal does not warrant good performance on real speech. Moreover, this stationary signal could be seen by the AEC as background noise; if so, the AEC would not operate during tests as it would do normally on real speech signals.

b) Use of a particular sequence of a non-stationary signal (e.g. artificial speech P.50, composite source signal[5)], real speech). This choice would eliminate the drawbacks of the former solution; however, since the chosen sequence would be a particular one, it is not warranted that the algorithms complying to the required performance would operate satisfactorily, on the average, on real speech. Moreover, such sequences have not yet been clearly identified and labelled. Note that sequences of non-stationary signals containing reproducible characteristics repeated several times during a typical measurement period (e.g. ≥ 1 s as presently specified in this Recommendation) like CSS are eventually better suited than artificial speech P.50 or real speech for convergence tests for example. This matter is under study.

c) Use of several (e.g. 10-20) sequences of a non-stationary signal (e.g. artificial speech P.50) chosen randomly (or from systematic evaluation) and averaging of the measurement results to estimate the performance. Although this choice would eventually be the most satisfactory one on a theoretical basis, the measurement procedure would take significantly more time than the measurements with a single sequence. The specification of sequences, averaging procedure and confidence intervals are under study.

Since the first solution a) may have several drawbacks, the two last solutions [b) and c)] are preferred. Using two different figures for the measurement of transient characteristics (e.g. Tic) and steady state characteristics (e.g. TCLwst) is under study.

5.2.2.2 Signal levels

The tests described in the sequel apply to level signals between [–30 dBm0] and [–10 dBm0] at R_{in}, and acoustic signals with levels –4.7 dBPa [$\pm$ 10 dBPa] at S_{in}.

5.2.3 Interface conditions – user side

5.2.3.1 Acoustic echo path

The use of real rooms or enclosures with appropriate acoustic characteristics is recommended. Echo paths simulated by electronic devices like digital reverberators with non time-varying reflection patterns can be used as well if the terminal has internal accesses on the user side. In this latter case, the electronic simulator adjustments should comply with the values recommended for real rooms or enclosures; moreover, the shape of the simulated impulse response envelope should be similar to the real echo path impulse response.

– For teleconference systems, the reverberation time averaged over the transmission bandwidth shall be typically 400 ms; the reverberation time in the lowest octave shall be no more than twice this average value; the reverberation time in the highest octave shall be not less than half this value. The volume of a typical test room shall be of the order of 90m^3 / 2700 ft^3.

[5)] The composite source signal (CSS) is described in the Supplement No. 21 to Recommendation P.34.

– For hands-free telephones and videophones, the reverberation time averaged over the transmission bandwidth shall be typically 500 ms; the reverberation time in the lowest octave shall be no more than twice this average value; the reverberation time in the highest octave shall be not less than half this value. The volume of a typical test room shall be of the order of 50 m^3 / 1500 ft^3.
– For mobile radio telephones an enclosure simulating the interior of a car can be used; a real car can be used as well. A typical average "reverberation time" is 60 ms. The volume of the enclosure shall be of the order of 2.5 m^3 / 75 ft^3.

5.2.3.2 Acoustic noise

Typical noise conditions are given in Recommendation P.34 (for measurement of hands-free telephones), and in the Supplements No. 15 and No. 16 to the P-Series Recommendations.

It is recommended to perform measurements in 1/3 octave bands for frequency-dependent quantities. Subjective effects like the spectral masking of echo by noise are under study.

5.2.3.3 Signals for double talk conditions

The general problem of double talk measurements is under study. The use of the composite source signal[6)] can be considered.

5.3 Correspondence between performance values and transmission delays

The performance values specified below correspond to the largest transmission delays (typically more than 250 ms) that can be encountered. In the case when the acoustic echo controller is aware of the existence of a shorter transmission delay (this may be possible for example in ISDN connection), those values can be somewhat relaxed. This point is under study.

5.4 Specifications and verification tests

5.4.1 Weighted terminal coupling loss – single talk (TCLwst)

Test procedure:

Step 1: All the AEC functional units are initially reset and then enabled.

Step 2: A signal is applied at R_{in} for a sufficient time (to be defined, under study) so that the different functional units (in particular the acoustic echo canceller) reach their steady states. No other speech signal than the acoustic return from the loudspeaker(s) is applied to the microphone(s).

Step 3: Make an electrical measurement of the signal at S_{out}. The value TCLwst is the difference (in dB) between the signal level before the enabling of the AEC and the signal level at this step in the test.

Requirements:

For teleconference systems and for hands-free communication on both sides, TCLwst shall be at least [40 dB].

For hands-free telephones and videophones interworking with distant users connected to the PSTN, TCLwst shall be at least [45 dB].

For mobile radio systems, TCLwst shall be at least [45 dB] when no acoustic noise is added at the S_{in} interface. When typical acoustic noise is present (e.g. car noise), masking of echo by noise can be taken into account (the masking effect depends on the levels and spectra of the echo and of the noise). The specification of adequate values for TCLwst according to different noise characteristics is under study.

---

6) The composite source signal (CSS) is described in the Supplement No. 21 to Recommendation P.34.

5.4.2 Weighted terminal coupling loss – double talk (TCLwdt)

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After the echo loss has attained TCLwst, an acoustic signal simulating the local user's speech is applied at the S_{in} point for [2] seconds.

Step 3: The processing unit is frozen, and then the simulated local speech is removed.

Step 4: Make an electrical measurement of the signal at S_{out}. The value TCLwdt is the difference (in dB) between the signal level before the enabling of the AEC and the signal level at this step in the test.

Requirements:

For teleconference systems and for hands-free communication on both sides, TCLwdt shall be at least [25 dB].

For hands-free telephones and videophones interworking with distant users connected to the PSTN, TCLwdt shall be at least [30 dB].

For mobile radio systems, TCLwdt shall be at least [30 dB] when no acoustic noise is added at the S_{in} interface. When typical acoustic noise is present (e.g. car noise), the same considerations as in 5.4.1 apply.

5.4.3 Received speech attenuation during double talk (Ardt)

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After the echo loss has attained TCLwst, a signal simulating the local user's speech is applied at the S_{in} point for [2] seconds.

Step 3: The processing unit is frozen, and then the simulated local speech is removed.

Step 4: Make an acoustical measurement of the signal at R_{out} with a signal at –20 dBm0 at R_{in}. The value Ardt is the difference in level at R_{out} between the measurement at this step in the test and the measurement at the end of step 1.

Requirements:

For all the applications, Ardt shall be no more than 6 dB.

5.4.4 Sent speech attenuation during double talk (Asdt)

Test procedure:

Step 1: The AEC is firstly operated as in the test of Ardt (steps 1 and 2).

Step 2: The processing unit is frozen, and then the signal at R_{in} is removed.

Step 3: Make an electrical measurement of the signal at S_{out} with a signal at –4.7 dBPa at S_{in}. The value Asdt is the difference in level at S_{out} between the measurement at this step in the test and the measurement when the terminal is reset.

Requirements:

For all the applications, Asdt shall be no more than 6 dB.

5.4.5 Received speech distortion during double talk (Drdt)

Test procedure:

Step 1: The processing unit is bypassed and the reference distortion measurement is made at R_{out} (acoustic signal).

Step 2: The AEC is operated as in the test of Ardt (steps 1 and 2).

Step 3: The processing unit is frozen, and then the simulated local speech at S_{in} is removed.

Step 4: Make a distortion measurement of the signal at R_{out}. The result of the measurement is called Drdt.

Requirements:

For all the applications, the supplementary distortion at R_{out} in comparison with single talk conditions should be low. The values of Drdt are under study.

5.4.6 Sent speech distortion during double talk (Dsdt)

Test procedure:

Step 1: The processing unit is bypassed and the reference distortion measurement is made at S_{out} (electric signal).

Step 2: The AEC is operated as in the test of Ardt (steps 1 and 2).

Step 3: The processing unit is frozen, and then the signal at R_{in} is removed.

Step 4: Make a distortion measurement of the signal at S_{out}. The result of the measurement is called Dsdt.

Requirements:

For all the applications, the supplementary distortion at S_{out} in comparison with single talk conditions should be low. The values of Dsdt are under study.

5.4.7 Maximum frequency shift (or pitch ratio) (Pr)

The test procedure is under study.

Requirements:

For all frequencies in the transmission band above 170 Hz the maximum frequency shift on either side shall be [3%]. The maximum absolute frequency shift shall be [5 Hz] for all frequencies in the transmission band below 170 Hz. Shifting up (i.e.increasing the pitch) is recommended for all 4 wires terminals.

It is recalled that frequency shift should be avoided in the case of terminals likely to be used on connections including network electric echo cancellers.

5.4.8 Break-in time – simple talk (Tonst)

5.4.8.1 Receive side

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After TCLwst has attained its recommended value, the signal applied at R_{in} is cut off and a signal simulating the local user's speech is applied at the S_{in} point for [2] seconds.

Step 3: The simulated local speech at S_{in} is cut off, the received signal is again applied at R_{in} and a timer is started.

Step 4: The timer is stopped when the attenuation of the signal at R_{out} has become lower than [3 dB]. The time interval measured is called Tonst_r.

Requirements:

For all the applications, Tonst_r shall be no more than [20 ms].

5.4.8.2 Send side

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After TCLwst has attained its recommended value, the signal applied at R_{in} is cut off, a signal simulating the local user's speech is applied at the S_{in} point and a timer is started.

Step 3: The timer is stopped when the attenuation of the signal at S_{out} has become lower than [3 dB]. The time interval measured is called Tonst_s.

Requirements:

For all the applications, Tonst_s shall be no more than [20 ms].

5.4.9 Break-in time – double talk (Tondt)

5.4.9.1 Receive side

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After TCLwst has attained its recommended value, the signal applied at R_{in} is cut off and a signal simulating the local user's speech is applied at the S_{in} point for [2] seconds.

Step 3: The received signal is again applied at R_{in} and a timer is started.

Step 4: After [20 ms] the timer is stopped, the processing unit is frozen and the signal at S_{in} is cut off.

Step 5: The acoustic signal level at R_{out} is measured. The time interval specified in step 4 is called Tondt_r.

Requirements:

For all the applications, the attenuation of the signal at R_{out} should be no more than 6 dB after [20 ms].

5.4.9.2 Send side

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After TCLwst has attained its recommended value, a signal simulating the local user's speech is applied at the S_{in} point and a timer is started.

Step 3: After [20 ms] the timer is stopped, the processing unit is frozen and the signal at R_{in} is cut off.

Step 4: The electric signal level at S_{out} is measured. The time interval specified in step 3 is called Tondt_s.

Requirements:

For all the applications, the attenuation of the signal at S_{out} should be no more than 6 dB after [20 ms].

5.4.10 Initial convergence time (Tic)

Test procedure:

Step 1: All the AEC functional units are initially reset and then enabled.

Step 2: A signal is applied at R_{in} and a timer is started.

Step 3: After [1] second, the processing unit is frozen.

Step 4: Make an electrical measurement of the signal at S_{out}. The time interval specified in step 3 is called Tic.

Requirements:

For all the applications, the attenuation of the echo shall be at least [20 dB] after Tic = [1] second.

5.4.11 Recovery time after double talk (Trdt)

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After TCLwst has attained its recommended value, the signal applied at R_{in} is cut off and a signal simulating the local user's speech is applied at the S_{in} point for [2] seconds.

Step 3: The received signal is again applied at R_{in}, and after [2] seconds the signal simulating the local user's speech is cut off; then a timer is started.

Step 4: After [1] second the timer is stopped and the processing unit is frozen.

Step 5: The electric signal level at S_{out} is measured. The time interval specified in step 4 is called Trdt.

Requirements:

For all the applications, the attenuation of the signal at S_{out} should be at least [20 dB] after Trdt = [1] second.

5.4.12 Terminal coupling loss during echo path variation (TCLwpv)

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After TCLwst has attained its recommended value, a simulated or real echo path variation is applied for [5] seconds (means to produce echo path variations are under study).

Step 3: At the end of the echo path variation, the processing unit is frozen, and the signal level at S_{out} is measured. The value TCLwpv is the difference (in dB) between the signal level before the enabling of the AEC and the signal level at this step in the test.

Requirements:

For all the applications, TCLwpv should be at least [10] dB.

5.4.13 Recovery time after echo path variation (Trpv)

Test procedure:

Step 1: The AEC is firstly operated as in the test of TCLwst (steps 1 and 2).

Step 2: After TCLwst has attained its recommended value, a simulated or real echo path variation is applied during [5] seconds (means to produce echo path variations are under study).

Step 3: At the end of the echo path variation a timer is started.

Step 4: After [1] second, the processing unit is frozen and the signal level at S_{out} is measured. The time interval specified in this step of the test is called Trpv.

Requirements:

For all the applications, the attenuation of the echo should be at least [20 dB] after Trpv = [1] second.

6 Specifications for interworking with the network

6.1 Scope

These specifications are the requirements for AECs to interwork properly with other devices on the network side which can be installed either in the terminal itself or in remote places in the network. They cover all the applications mentioned in 1.2.

6.2 Interworking with speech codecs

There are two cases of interest here: speech coding on the R_{in} path and speech coding on the S_{out} path.

6.2.1 R_{in} path speech coding

6.2.1.1 Bit rate

The obvious item is that the bit rate at the edge of the processing unit including the AEC at R_{in} must match the incoming rate.

6.2.1.2 Bandwidth

The other consideration is that bandwidth of the AEC must, if at all possible, match the incoming bandwidth within a reasonable tolerance (see 4.4). If the AEC is not capable of equalling the incoming bandwidth, it must introduce compensating filters.

6.2.1.3 Tandem speech coding effects

If, in the processing inside the AEC, the speech is re-encoded, the effects of tandeming with acceptable R_{in} speech codings must be considered. As a side note, the frequency shift method of howling control could be considered in this category. Also, any speech activity detection which results in processing of the R_{in} signal before being played out on R_{out} would be considered in this category. It is under study whether d.c. removal should also be in this category – current thinking is no. Overall, these tandem effects are considered to be less serious than those applying to the S_{in} to S_{out} path (6.2.2.3) since any effect in this section would be localized to the user of the AEC.

6.2.2 S_{out} path speech coding

6.2.2.1 Bit rate

The obvious item is that the bit rate at the edge of the processing unit including the AEC at S_{out} must match the outgoing line rate.

6.2.2.2 Bandwidth

The other consideration is that bandwidth of the AEC must, if at all possible, match the outgoing bandwidth within a reasonable tolerance (see 4.4). If the AEC is not capable of equalling the outgoing bandwidth, it must introduce compensating filters.

6.2.2.3 Tandem speech coding effects

If, in the processing inside the AEC, the speech is re-encoded, the effects of tandeming with acceptable S_{out} speech codings must be considered. The echo cancelling process itself is a form of speech coding in the sense of this section and could have implications for the external speech coder on S_{out}. In general, processing which is representable as a linear filter would not pose a problem but processing which is non-linear might pose a problem. As a side note, the frequency shift method of howling control could be considered in this category. Also, any speech activity detection which results in modulating the S_{in} signal before being played out on S_{out} would be considered in this category. It is under study whether d.c. removal should also be in this category – current thinking is no. These tandem effects are considered to be more serious than the ones applying to the R_{in} to R_{out} path (6.2.1.3) since these effects are observable to the rest of the network and to the (possibly) non-AEC user. It is particularly important that the AEC not negatively impact the performance of cascaded network equipment.

6.3 Interworking with network echo cancellers

The most important item, from the point of view of network echo cancellers, is that there be no signal (or signal component) on S_{out} which appears to be correlated with R_{in} (within the constraints of the tail delay of the network echo canceller) unless it is during an instance of obvious double-talk. For purposes of discussion, obvious double-talk is defined as having speech power at S_{out} greater than the power at R_{in}. In fact, this constraint is perhaps a bit generous since one might (incorrectly) read it to permit non-linear distortions in the echo path from R_{in} to S_{out}. A safer (but perhaps more strenuous than is required) condition would be to require that after initial convergence of the AEC the level at S_{out} be more than [40 dB] down, including non-linear processing, from the level at R_{in} unless double-talk is occurring. This latter condition protects the network echo canceller since there is no echo on which to adapt (see Figure 3).

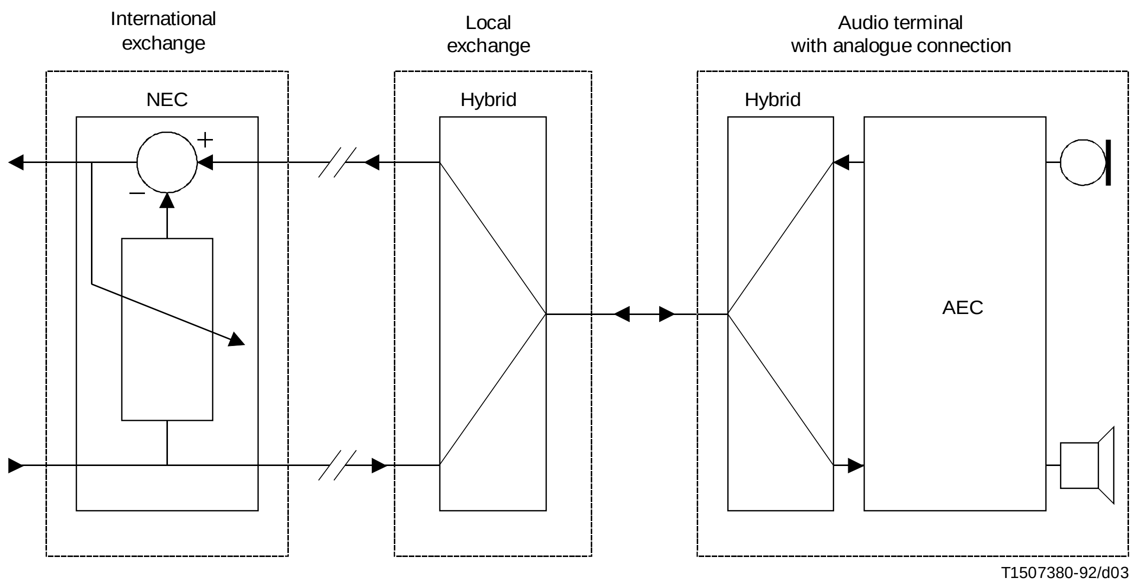

FIGURE 3/G.167

Example of network configuration where disturbing interaction can occur between acoustic echo controller (AEC) and network echo canceller (NEC)

6.4 DCME and PCME interactions

The constraints mentioned above for speech coding and network echo cancellers should largely satisfy the needs of DCME and PCME equipment. One additional factor that is of interest is background noise effects. The AEC must keep the background noise level on S_{out} as constant and as low as possible. If the AEC includes a non-linear device like a center-clipper or if it introduces a large amount of loss on the send path when there is speech on the receive side only, it is recommended that appropriate comfort noise be inserted by the AEC to avoid false activation of the DCME/PCME speech detectors on modulated background noise which can result from non-linear processing or loss insertion.

6.5 Interworking between a wide-band terminal and other types of terminals through the network

A mechanism has been defined in Recommendation G.725 which allows interoperability between different types of terminals according to different transmission modes. After initialization of the communication the local and distant terminals are in a common mode. The AEC in the local terminal shall operate according to the pass-band and the reference levels corresponding to that mode. In case of call transfer, care must be taken for proper reinitialization of the AEC (if changing from one mode to another mode) to avoid instability and reduce transient effects as much as possible (under study).