

166166



G00002

Process for assessing the transmission quality of a speech transmission line

The invention relates to a process for assessing the transmission quality of a speech transmission line according to the preamble of claim 1.

Providing a good transmission quality is the aim of all network suppliers, as long as there are electric transmission lines. The question what has to be understood by quality and how it is to be detected by the measuring technique is, however, still a considerable problem. A first method for

testing a transmission quality is known from the document GB-1 455 712. There is a comparison between an input test signal and the corresponding output signal. The difference signal serves as a value for the rating of a quantity "quality" that is not defined in details.

Document EP-A-0 114 463 describes an arrangement for measuring the connection quality between a transmitter and a receiver. For this purpose, a multitude of signal tones is continuously transmitted. The received tones are digitized and transformed by means of a Fourier transformation from the time domain into the frequency domain. Special parameters of the transformed tones serve for rating the connection quality.

A further method for measuring the quality is known from the document DE-A-35 13 361, in particular for radio transmission channels during the on-going transmission. For detecting the co-channel interferences, fast and slow variations are measured separately and from those a quality indicator is derived.

It is common to all methods mentioned above that more or less arbitrary indications for rating the quantity "quality" are used. These indications are in no direct relation to the subjective human perception, in particular for the assessment, whether human speech can be clearly understood by a receiving person or not. For this reason, it is usual nowadays to order trained test persons to critical reception points for e. g. mobile telephone connections. These persons decide on the basis of a spoken text, e. g. selected sentences in an arbitrary language, whether the actual reception qualities are satisfactory or not sufficient and therefore unsatisfactory.

It is evident that the test method with test persons is very expensive and time consuming. Furthermore, this method succumbs to subjective variations, which go back to differences in the different test persons, their actual condition and their moods.

It is therefore the object of the invention to provide a method that makes it possible to automatically gain reproducible assessment results that correspond more or less to the results that the mentioned test persons would determine. This means that a method has to be indicated that automatically allows a quality assessment imitating the human auditory sensation.

The solution of this object is given in the characterizing part of claim 1. The depending claims give embodiments of the invention.

In the following the invention is described by means of the example shown in the only figure. The figure shows the block circuit of a transmission line for the case of a speech connection between a wired telephone set 11 and a mobile telephone 12, e. g. a car telephone. The stationary telephone set 11 belongs e. g. to a private network 13, which is symbolized by an inhouse telephone exchange 14. This inhouse telephone exchange 14 is connected with the public telephone network 16 by means of a stationary installed connection line 15, the public telephone network being symbolized by a public telephone exchange 17.

The public telephone network 16 cooperates now with a mobile telephone network 18, what is symbolized by a stationary mobile node 19 with antenna 20 and by a connection line 21. Between the antenna 20 and the mobile telephone 12 there is

finally a radio communication link 22, which is generally speaking most critical with respect to the transmission quality between the two telephones 11, 12.

For assessing the transmission quality a measuring apparatus to be described later on is connected in such a way to the two terminals of the transmission line that it uses the transmission line in the same way as or at least very similar to a human user. This means that the mentioned terminals are arranged as close as possible to the microphone capsule of the transmitting telephone and the telephone receiver of a telephone at the receiving end, respectively. For practical reasons and because the influences of the terminal units, i. e. the telephones are very small and known, it is of course possible that unproblematic interfaces are used as "terminals", e. g. a telephone wall plug at the end of the network and the radio communication link 22, as long as the circuits of the measuring apparatus belonging to the antenna are identical with those of a mobile telephone 12.

In the figure, reference numeral 31 designates a mobile measuring unit that is e. g. installed in an automobile and is powered by the battery of the automobile. This measuring unit 31 is connected with the mobile node 19 via a radio communication link 22', which is to be tested. In addition, there is a stationary measuring unit 32, which is connected with the telephone network 13. The measuring units 31 and 32 form together the mentioned measuring apparatus, by means of which the quality of the transmission line via the private network 13, the public network 16 and in particular the radio communication link 22, 22', respectively, can be tested.

The process for assessing the transmission quality is based on the fact that the sequences SS of human speech are used as test signals. In this the process coincides with the processes mentioned at the beginning, where human test persons perform the assessment. The use of a short sentence has been proved to be particularly favorable, whereby the parts of the sentence are alternately spoken by a female and a male voice. The duration of the sentence should be between about 2 to 10 seconds, e. g. 5 seconds. Longer sequences are easily generated by repeating the short sentence an arbitrary number of times and immediately transmitted in a sequence and received in the same way. The language, in which the sentence is spoken, does not matter. It is therefore possible to use any language.

The process furthermore is based on the characteristics of neuronal networks, in particular on the trainability of such networks. Furthermore it is based on suitable pre- and post-processing methods, that reduce and change the data to be processed in such a way that input and output values are generated that are appropriate for the applied neuronal network.

A neuronal network is meant to be an arrangement that connects a multitude of input neurons with a very limited number of output neurons via an (other) multitude of connections and via a covered layer of neurons. This allows after a training phase to recognize arbitrary predefined output patterns. Reference is made at this point e. g. to S. Grossberg, "Nonlinear Neuronal Networks: Principles, Mechanisms and Architectures", Neuronal Networks Nr. 1, 1988, S. 17-61.

As mentioned, a preprocessing unit is connected ahead and optionally a postprocessing unit is connected behind the

neuronal network. The preprocessing unit receives via its input the speech sequences SS to be assessed, which have gone through the transmission line to be tested. The preprocessing unit converts these sequences SS into intermediate values ZW, which flow into the neuronal network. The neuronal network assesses the respective intermediate values ZW and assigns to those and the corresponding speech sequence SS via its different output activations a number of e. g. two or three possible assessment values WW. The possible assessment values can be assigned to the terms "good" and "bad" or "good", "medium" and "bad", which indicate in each case the transmission quality in a way understandable by man.

In particular, at least the following process steps are provided for the mentioned preprocessing:

- A synchronization pattern SM is positioned ahead of the speech sequence SS to be used, the synchronization pattern being easy recognizable at the receiving end by a cross-correlation. This pattern SM serves to fix a start time as exact as possible for each received sequence SS and has a time duration of about 50 to 100 ms.

- At the receiving end, the speech sequence SS has to be available in a digitized form. In a digital transmission, digital samples or amplitude values AP are available in standardized time intervals as a matter of fact. In an analogue transmission the analogue receiving signal has to be sampled in a known way, which yields in predetermined sampling steps AS of e. g. 125 µs corresponding to 8 kbit/s one amplitude value AP each.

- As soon as the starting time is determined, the following speech sequence SS is divided into a plurality of time intervals ZI. The division may take place in such a way, that all time intervals have the same length. This means e. g. that the sequence SS is divided into 300 intervals ZI of 256 sampling steps AS. It is, however, also possible to make time intervals of unequal length, e. g. depending on the type of the actual speech signals in each case, i. e. depending on a speech pause, a vowel etc. Furthermore, the intervals ZI can overlap each other by e. g. 50%. This means that with 256 sampling steps AS per time interval, 128 of these sampling steps are equal with those of the next time interval ZI. This reduces the probability for very abrupt transitions between neighboring intervals ZI.

- For each time interval ZI of the speech sequence SS the characteristic frequency spectrum of the different amplitude values AP of all sampling steps AS is determined, the frequency spectrum being assigned to the corresponding time interval ZI.

- The respective frequency spectrum is assessed with respect to the energy in it and is assigned to one of e. g. 18 classes and is normalized.

- Difference values of all normalized values of all time intervals ZI of the complete speech sequence SS with respect to the corresponding values of a reference sequence that has been received at optimum transmission quality are generated. By this, the deviation of the respective speech sequence SS to be assessed from the best expectable conditions are gained with optimal selectivity. Altogether, e. g. 5400 difference values are generated. These differ-

ence values form essentially the mentioned intermediate values ZW, i. e. the input values for the neuronal network.

An arbitrary number of additional steps can be added to the mentioned process steps of the preprocessing. Such additional steps are:

- Filtering by e. g. high pass filter and selective amplification respectively for relatively rising the higher frequencies.

- Forming the values of a time interval ZI by means of a Hamming window for eliminating extreme boundary values.

- Normalizing the amplitude values in order to take different volume levels of the received speech sequence SS into account.

The mentioned optional postprocessing of the output activations and of the assessment values WW of the neuronal network respectively consists in an interpretation of these values WW by a Fuzzy-Logic. A Fuzzy-Logic is able to determine in each case an assigned probability value based on the type of the corresponding activations at the output of the neuronal network. By this, the number of the possible probability values becomes greater than the number of the possible assessment values WW. Thereby, the human understandable assessments "very good", "good", "satisfactory", "sufficient" and "unsatisfactory" can be assigned to the probability values.

It is now important to train the assessment logic, which is necessary for the decisive assessment step, in a preparatory

phase. For this purpose, the predetermined speech sequence SS is assessed by suitable persons with respect to the quality and speech understandability at all conceivable transmission qualities at the receiving end. The different received speech sequences SS and the assigned human assessments are then commonly fed to the measuring unit 31 and/or 32 at the receiving end. By this the corresponding measuring unit and the neuronal network contained in it respectively, is learning to assess the speech sequences SS of any quality on its own, i. e. to assign the mentioned assessments "good", "medium" or "bad".

After the preparatory phase the measuring unit 31, 32 at the receiving end is acting during the actual working phase, i. e. during a second phase of arbitrary duration, with respect to the speech assessment in the same way as or at least similar to a human test person.

All mentioned process steps can be implemented in combination with the neuronal network on a computer. The two measuring units 31, 32 are therefore according to a preferred embodiment computers, e. g. personal computers. The processing and memory power can be selectively enhanced e. g. by an appropriate additional board. Furthermore, the speech sequence SS to be transmitted can also be stored digitally. Such personal computers can be used as measuring units at the transmitting end as well as at the receiving end. By this it is easily possible to perform the assessment of the transmission quality in both directions, either successively or simultaneously.

The computer of the two measuring units 31, 32 are furthermore capable to store the determined assessment and output values AW in an appropriate way. This means e. g. the

storage of the data together with time and place indications allowing evalutations of any type later on.

It is mainly a practical question how the described measuring units 31, 32, 33 are built up. Preferably, a complete assessment possibility is provided for each transmission direction, i. e. from the mobile measuring unit 31 to the stationary measuring unit 32 and vice versa from the stationary measuring unit 32 to the mobile measuring unit 31. This means that both measuring units 31, 32 are made for transmitting the speech sequences SS and for receiving the sequences transmitted from the other measuring unit. Furthermore, the assessement of the transmission quality from the mobile measuring unit 31 to another mobile measuring unit is possible.

As the above description shows, there is a considerable plurality of variations for performing the described process. Accordingly e. g. the neuronal network can be constructed in such a way, that five possible assessment values WW are outputed instead of the mentioned two ore three values. Furthermore, the way of the activations at the outputs can vary. Decisive for the variations is in any case, that speech sequences SS serve for the assessment of the transmission quality, the speech sequences being assessed by a trained assessment logic with respect to understandability at the receiving place in the same or at least in a similar way as by a trained test person, that means according to subjective human criteria.

CLAIMS

1. Process for assessing the transmission quality of a speech transmission line, wherein test signals are transmitted at the transmitting end and the corresponding receiving signals at the receiving end are assessed with respect to transmission quality, characterized in

   - that each test signal is a predetermined sequence (SS) of the human language, which is headed by a synchronization pattern (SM),

   - that at least three serial processing steps are provided at the receiving end,

   - that in the first processing step the respective synchronization pattern (SM) is recognized and by this the start time of the following sequence (SS) is determined,

   - that in the second processing step a first plurality of corresponding intermediate values (ZW) is formed by several serial substeps, wherein the respective sequence (SS) is divided into a first plurality of time intervals (ZI), which comprises in each case a second plurality of discrete, chronological amplitude values (AP), whereby in each case the appropriate class of a third plurality of classes is assigned to the corresponding amplitude values (AP) of each time interval (ZI), and wherein for each time interval (ZI) the difference between the corresponding class and that class is generated, which is assigned to the corresponding

time interval (ZI) for a sequence (SS) received at optimum transmission quality, whereby each difference corresponds to an intermediate value (ZW), and

- that in the third process step a trained neuronal network assignes the entire set of all intermediate values (ZW) to the corresponding assessment value (WW) of a plurality of possible assessment values, whereby the assessment value (WW) corresponds to the respective transmission quality.

2. Process according to claim 1, characterized in that the sequence (SS) has a length of about 2 to 10 seconds.

3. Process according to claim 1, characterized in that the sequence (SS) is a sentence of an arbitrary language, which is partially spoken by a male and partially by a female voice.

4. Process according to claim 1, characterized in that the sequence (SS) is repeated an arbitrary number of times and is used in immediate succession.

5. Process according to claim 1, characterized in that the synchronization pattern (SM) has a length of about 50 to 100 ms and is formed in such a way, that it is easily recognizable by means of cross-correlation.

6. Process according to claim 1, characterized in that the sequence (SS) is divided in a first plurality of time intervals (ZI) of equal length.

7. Process according to claim 1, characterized in that the sequence (SS) is divided in a first plurality of time intervals (ZI) of unequal length.

8. Process according to claim 1, characterized in that the neuronal network is trained before its actual use in such a way that sequences (SS) of the receiving end corresponding to all possible different transmission qualities and in combination with that the assigned human assessment of these transmission qualities are repeatedly inputed.

9. Process according to claim 1, characterized in that the first plurality is 300, in that the second plurality is 256, in that the third plurality is 18 and in that the plurality of possible assessment values is between 2 and 5.

10. Process according to claim 1, characterized in that a fourth processing step is provided by a Fuzzy-Logic, the fourth step enlarging the plurality of possible assessment values (WW) of the neuronal network.

11. Device for assessing the transmission quality of a speech transmission line, comprising means for sending test signals at a transmitting end and means for assessing the corresponding receiving signals with respect to transmission quality at the receiving end, characterized in

    a. that the means at the transmitting end are built such that each test signal is a predetermined sequence (SS) of the human language, which is headed by a synchronization pattern (SM),

    b. that the means at the receiving end are built such that at least three serial processing steps are provided such

    c. that in the first processing step the respective synchronization pattern (SM) is recognized and by this the start time of the following sequence (SS) is determined,

    d. that in the second processing step a first plurality of corresponding intermediate values (ZW) is formed by several serial substeps, wherein the respective sequence (SS) is divided into a first plurality of time intervals (ZI), which comprises in each case a second plurality of discrete, chronological amplitude values (AP), whereby in each case the appropriate class of a third plurality of classes is assigned to the corresponding amplitude values (AP) of each time interval (ZI), and wherein for each time interval (ZI) the difference between the corresponding class and that class is generated, which is assigned to the corresponding time interval (ZI) for a sequence (SS) received at optimum transmission quality, whereby each difference corresponds to an intermediate value (ZW), and

    e. that in the third processing step a trained neuronal network assignes an assessment value (WW) of a plurality of possible assessment values to the entire set of all intermediate values (ZW), whereby the assessment value (WW) corresponds to the respective transmission quality.

12. Device according to claim 11, characterized in that the means at the transmitting end are built such that the sequence (SS) has a length of about 2 to 10 seconds.

13. Device according to claim 11, characterized in that the means at the transmitting end are built such that the sequence (SS) is a sentence of an arbitrary language, which is partially spoken by a male and partially by a female voice.

14. Device according to claim 11, characterized in that the means at the transmitting end are built such that the sequence (SS) is repeated an arbitrary number of times and is used in immediate succession.

15. Device according to claim 11, characterized in that the means at the transmitting end are built such that the synchronization pattern (SM) of the test signal has a length of about 50 to 100 ms and is formed in such a way, that it is easily recognizable by means of cross-correlation.

16. Device according to claim 11, characterized in that the means at the receiving end are built such that the sequence (SS) is divided in a first plurality of time intervals (ZI) of equal length.

17. Device according to claim 11, characterized in that the means at the receiving end are built such that the sequence (SS) is divided in a first plurality of time intervals (ZI) of unequal length.

8. Device according to claim 11, characterized in that the neuronal network in the means at the receiving end is built such that it has been trained before its actual use in such a way that sequences (SS) of the receiving end corresponding to all possible different transmission qualities and in combination with that the assigned human assessment of these transmission qualities are repeatedly inputed.

9. Device according to claim 11, characterized in that the means at the receiving end are built such that the first plurality is 300, that the second plurality is 256, that the third plurality is 18 and that the plurality of possible assessment values is between 2 and 5.

20. Device according to claim 11, characterized in that a Fuzzy-Logic for a fourth processing step is provided wherein the plurality of possible assessment values (WW) of the neuronal network is enlarged.

21. Computer program product for carrying out a process for assessing the transmission quality of a speech transmission line, wherein test signals are transmitted at the transmitting end and the corresponding receiving signals at the receiving end are assessed with respect to transmission quality, characterized in

a. that each test signal is a predetermined sequence (SS) of the human language, which is headed by a synchronization pattern (SM),

b. that at least three serial processing steps are provided at the receiving end,

c. that in the first processing step the respective synchronization pattern (SM) is recognized and by this the start time of the following sequence (SS) is determined,

d. that in the second processing step a first plurality of corresponding intermediate values (ZW) is formed by several serial substeps, wherein the respective sequence (SS) is divided into a first plurality of time intervals (ZI), which comprises in each case a second plurality of discrete, chronological amplitude values (AP), whereby in each case the appropriate class of a third plurality of classes is assigned to the corresponding amplitude values (AP) of each time interval (ZI), and wherein for each time interval (ZI) the difference between the corresponding class and that class is generated, which is assigned to the corresponding time interval (ZI) for a sequence (SS) received at optimum transmission quality, whereby each difference corresponds to an intermediate value (ZW), and

e. that in the third processing step a trained neuronal network assignes the entire set of all intermediate values (ZW) to the corresponding assessment value (WW) of a plurality of possible assessment values, whereby the assessment value (WW) corresponds to the respective transmission quality.

2. Computer program product according to claim 21, characterized in that the sequence (SS) has a length of about 2 to 10 seconds.

23. Computer program product according to claim 21, characterized in that the sequence (SS) is a sentence of an arbitrary language, which is partially spoken by a male and partially by a female voice.

24. Computer program product according to claim 21, characterized in that the sequence (SS) is repeated an arbitrary number of times and is used in immediate succession.

25. Computer program product according to claim 21, characterized in that the synchronization pattern (SM) of the test signal has a length of about 50 to 100 ms and is formed in such a way, that it is easily recognizable by means of cross-correlation.

26. Computer program product according to claim 21, characterized in that the sequence (SS) is divided in a first plurality of time intervals (ZI) of equal length.

27. Computer program product according to claim 21, characterized in that the sequence (SS) is divided in a first plurality of time intervals (ZI) of unequal length.

28. Computer program product according to claim 21, characterized in that the neuronal network is trained before its actual use in such a way that sequences (SS) of the receiving end corresponding to all possible different transmission qualities and in combination with that the assigned human assessment of these transmission qualities are repeatedly inputed.

29. Computer program product according to claim 21, characterized in that the first plurality is 300, in that the second plurality is 256, in that the third plurality is 18 and in that the plurality of possible assessment values is between 2 and 5.

30. Computer program product according to claim 21, characterized in that a fourth processing step is provided by a Fuzzy-Logic, the fourth processing step enlarging the plurality of possible assessment values (WW) of the neuronal network.