In embodiments the quality measure represents a perceptual quality measure. A single value which is an estimation of the subjective quality of the first coding algorithm and a single value which is an estimation of the subjective quality of the second coding algorithm may be computed. The encoding algorithm which gives the best estimated subjective quality may be chosen just based on the comparison of these two values. This is different from what is done in the AMR-WB+ standard where many features representing different characteristics of the signal are computed and, then, a classifier is applied to decide which algorithm to choose.

In embodiments, the respective quality measure is estimated based on a portion of the weighted audio signal, i.e. a weighted version of the audio signal. In embodiments, the weighted audio signal can be defined as an audio signal filtered by a weighting function, where the weighting function is a weighted LPC filter A(z/g) with A(z) an LPC filter and g a weight between 0 and 1 such as 0.68. It turned out that good measures of perceptual quality can be obtained in this manner. Note that the LPC filter A(z) and the weighted LPC filter A(z/g) are determined in a pre-processing stage and that they are also used in both encoding algorithms. In other embodiments, the weighting function may be a linear filter, a FIR filter or a linear prediction filter.

In embodiments, the quality measure is the segmental SNR (signal to noise ratio) in the weighted signal domain. It turned out that the segmental SNR in the weighted signal domain represents a good measure of the perceptual quality and, therefore, can be used as the quality measure in a beneficial manner. This is also the quality measure used in both ACELP and TCX encoding algorithms to estimate the encoding parameters.

Another quality measure may be the SNR in the weighted signal domain. Other quality measures may be the segmental SNR, the SNR of the corresponding portion of the audio signal in the non-weighted signal domain, i.e. not filtered by the (weighted) LPC coefficients. Other quality measures may be the cepstral distortion or the noise-to-mask ratio (NMR).

Generally, SNR compares the original and processed audio signals (such as speech signals) sample by sample. Its goal is to measure the distortion of waveform coders that reproduce the input waveform. SNR may be calculated as shown in Fig. ~~5a~~4a, where x(i) and y(i) are the original and the processed samples indexed by i and N is the total number

of samples. Segmental SNR, instead of working on the whole signal, calculates the average of the SNR values of short segments, such as 1 to 10 ms, such as 5ms. SNR may be calculated as shown in Fig. ~~5b~~4b, where N and M are the segment length and the number of segments, respectively.

In embodiments of the invention, the portion of the audio signal represents a frame of the audio signal which is obtained by windowing the audio signal and selection of an appropriate encoding algorithm is performed for a plurality of successive frames obtained by windowing an audio signal. In the following specification, in connection with the audio signal, the terms “portion” and “frame” are used in an exchangeable manner. In embodiments, each frame is divided into subframes and segmental SNR is estimated for each frame by calculating SNR for each subframe, converted in dB and calculating the average of the subframe SNRs in dB.

Thus, in embodiments, it is not the (segmental) SNR between the input audio signal and the decoded audio signal that is estimated, but the (segmental) SNR between the weighted input audio signal and the weighted decoded audio signal is estimated. As far as this (segmental) SNR is concerned, reference can be made to chapter 5.2.3 of the AMR-WB+ standard (International Standard 3GPP TS 26.290 V6.1.0 2004-12).

In embodiments of the invention, the respective quality measure is estimated based on the energy of a portion of the weighted audio signal and based on an estimated distortion introduced when encoding the signal portion by the respective algorithm, wherein the first and second estimators are configured to determine the estimated distortions dependent on the energy of a weighted audio signal.

In embodiments of the invention, an estimated quantizer distortion introduced by a quantizer used in the first encoding algorithm when quantizing the portion of the audio signal is determined and the first quality measure is determined based on the energy of the portion of the weighted audio signal and the estimated quantizer distortion. In such embodiments, a global gain for the portion of the audio signal may be estimated such that the portion of the audio signal would produce a given target bitrate when encoded with a quantizer and an entropy encoder used in the first encoding algorithm, wherein the estimated quantizer distortion is determined based on the estimated global gain. In such embodiments, the estimated quantizer distortion may be determined based on a power of the estimated gain. When the quantizer used in the first encoding algorithm is a uniform

Claims

1. Apparatus ~~(10)~~ for selecting one of a first encoding algorithm having a first characteristic and a second encoding algorithm having a second characteristic for encoding a portion of an audio signal ~~(40)~~ to obtain an encoded version of the portion of the audio signal ~~(40)~~, comprising:

a first estimator ~~(12)~~ for estimating a first quality measure for the portion of the audio signal, the first quality measure being associated with the first encoding algorithm, without actually encoding and decoding the portion of the audio signal using the first encoding algorithm;

a second estimator ~~(14)~~ for estimating a second quality measure for the portion of the audio signal, the second quality measure being associated with the second encoding algorithm, without actually encoding and decoding the portion of the audio signal using the second encoding algorithm; and

a controller ~~(16)~~ for selecting the first encoding algorithm or the second encoding algorithm based on a comparison between the first quality measure and the second quality measure,

wherein the first and second quality measures are SNRs (signal to noise ratio) or segmental SNRs of the corresponding portion of a weighted version of the audio signal.

2. Apparatus ~~(10)~~ of claim 1, wherein the first encoding algorithm is a transform coding algorithm, a MDCT (modified discrete cosine transform) based coding algorithm or a TCX (transform coding excitation) coding algorithm and wherein the second encoding algorithm is a CELP (code excited linear prediction) coding algorithm or an ACELP (algebraic code excited linear prediction) coding algorithm.

3. Apparatus ~~(10)~~ of one of claims 1 or 2, wherein the first estimator ~~(12)~~ is configured to determine an estimated quantizer distortion which a quantizer used in the first encoding algorithm would introduce when quantizing the portion of the audio signal and to estimate the first quality measure based on an energy of a

portion of a weighted version of the audio signal and the estimated quantizer distortion.

4. Apparatus ~~(10)~~ of claim 3, wherein the first estimator ~~(12)~~ is configured to estimate a global gain for the portion of the audio signal such that the portion of the audio signal would produce a given target bitrate when encoded with a quantizer and an entropy coder used in the first encoding algorithm, wherein the first estimator ~~(12)~~ is further configured to determine the estimated quantizer distortion based on a power of the estimated global gain, wherein the quantizer used in the first encoding algorithm is a uniform scalar quantizer and wherein the first estimator ~~(12)~~ is configured to determine the estimated quantizer distortion using the formula D = G*G/12, wherein D is the estimated quantizer distortion and G is the estimated global gain.

5. Apparatus ~~(10)~~ of one of claims 3 or 4, wherein the first quality measure is a segmental SNR of a portion of the weighted audio signal and wherein the first estimator ~~(12)~~ is configured to estimate the segmental SNR by calculating an estimated SNR associated with each of a plurality of sub-portions of the portion of the weighted audio signal based on an energy of the corresponding sub-portions of the weighted audio signal and the estimated quantizer distortion and by calculating an average of the SNRs associated with the sub-portions of the portion of the weighted audio signal to obtain the estimated segmental SNR for the portion of the weighted audio signal.

6. Apparatus ~~(10)~~ of one of claims 1 to 5, wherein the second estimator ~~(14)~~ is configured to determine an estimated adaptive codebook distortion which an adaptive codebook used in the second encoding algorithm would introduce when using the adaptive codebook to encode the portion of the audio signal, and wherein the second estimator ~~(14)~~ is configured to estimate the second quality measure based on an energy of a portion of a weighted version of the audio signal and the estimated adaptive codebook distortion, wherein, for each of a plurality of sub-portions of the portion of the audio signal, the second estimator ~~(14)~~ is configured to approximate the adaptive codebook based on a version of the sub-portion of the weighted audio signal shifted to the past by a pitch-lag determined in a pre-processing stage, to estimate an adaptive codebook gain such that an error between the sub-portion of the portion of the weighted audio signal and the

approximated adaptive codebook is minimized, and to determine the estimated adaptive codebook distortion based on the energy of an error between the sub-portion of the portion of the weighted audio signal and the approximated adaptive codebook scaled by the adaptive codebook gain.

7. Apparatus ~~(10)~~ of claim 6, wherein the second estimator ~~(14)~~ is further configured to reduce the estimated adaptive codebook distortion determined for each sub-portion of the portion of the audio signal by a constant factor.

8. Apparatus ~~(10)~~ of claim 6 or 7, wherein the second quality measure is a segmental SNR of the portion of the weighted audio signal, and wherein the second estimator ~~(14)~~ is configured to estimate the segmental SNR by calculating an estimated SNR associated with each sub-portion based on the energy of the corresponding sub-portion of the weighted audio signal and the estimated adaptive codebook distortion and by calculating an average of the SNRs associated with the sub-portions to obtain the estimated segmental SNR for the portion of the weighted audio signal.

9. Apparatus ~~(10)~~ of claim 6, wherein the second estimator ~~(14)~~ is configured to approximate the adaptive codebook based on a version of the portion of the weighted audio signal shifted to the past by a pitch-lag determined in a pre-processing stage, to estimate an adaptive codebook gain such that an error between the portion of the weighted audio signal and the approximated adaptive codebook is minimized, and to determine the estimated adaptive codebook distortion based on the energy of an error between the portion of the weighted audio signal and the approximated adaptive codebook scaled by the adaptive codebook gain.

10. Apparatus ~~(10)~~ of one of claims 1 to 9, wherein the controller ~~(16)~~ is configured to utilize a hysteresis in comparing the estimated quality measures.

11. Apparatus ~~(20)~~ for encoding a portion of an audio signal, comprising the apparatus ~~(10)~~ according to one of claims 1 to 10, a first encoder stage ~~(26)~~ for performing the first encoding algorithm and a second encoder stage ~~(28)~~ for performing the second encoding algorithm, wherein the apparatus for encoding ~~(20)~~ is configured

to encode the portion of the audio signal using the first encoding algorithm or the second encoding algorithm depending on the selection by the controller ~~(16)~~.

12. System for encoding and decoding comprising an apparatus ~~(20)~~ for encoding according to claim ~~18~~11 and a decoder configured to receive the encoded version of the portion of the audio signal and an indication of the algorithm used to encode the portion of the audio signal and to decode the encoded version of the portion of the audio signal using the indicated algorithm.

13. Method for selecting one of a first encoding algorithm having a first characteristic and a second encoding algorithm having a second characteristic for encoding a portion of an audio signal to obtain an encoded version of the portion of the audio signal, comprising:

estimating a first quality measure for the portion of the audio signal, the first quality measure being associated with the first encoding algorithm, without actually encoding and decoding the portion of the audio signal using the first encoding algorithm;

estimating a second quality measure for the portion of the audio signal, the second quality measure being associated with the second encoding algorithm, without actually encoding and decoding the portion of the audio signal using the second coding algorithm; and

selecting the first encoding algorithm or the second encoding algorithm based on a comparison between the first quality measure and the second quality measure,

wherein the first and second quality measures are SNRs (signal to noise ratio) or segmental SNRs of the corresponding portion of a weighted version of the audio signal.

14. Method of claim 13, wherein the first encoding algorithm is a transform coding algorithm, a MDCT (modified discrete cosine transform) based coding algorithm or a TCX (transform coding excitation) coding algorithm and wherein the second encoding algorithm is a CELP (code excited linear prediction) coding algorithm or an ACELP (algebraic code excited linear prediction) coding algorithm.

15. Method of one of claims 13 or 14, comprising determining ~~(108)~~ an estimated quantizer distortion which a quantizer used in the first coding algorithm would introduce when quantizing the portion of the audio signal and determining the quality measure based on an energy of a portion of a weighted version of the audio signal and the estimated quantizer distortion.

16. Method of claim 15, comprising estimating ~~(106)~~ a global gain for the portion of the audio signal such that the portion of the audio signal would produce a given target bitrate when encoded with a quantizer and an entropy coder used in the first coding algorithm, and determining ~~(108)~~ the estimated quantizer distortion based on a power of the estimated global gain, wherein the quantizer is a uniform scalar quantizer, wherein the estimated quantizer distortion is determined using the formula D = G*G/12, wherein D is the estimated quantizer distortion and G is the estimated global gain.

17. Method of one of claims 15 or 16, wherein the first quality measure is a segmental SNR of the LPC filtered version of a portion of the weighted audio signal, and comprising estimating the first segmented SNR by calculating an estimated SNR associated with each of a plurality of sub-portions of the portion of the weighted audio signal based on an energy of the corresponding sub-portions of the weighted audio signal and the estimated quantizer distortion and by calculating an average of the SNRs associated with the sub-portions of the portion of the weighted audio signal to obtain the estimated segmental SNR for the portion of the weighted audio signal.

18. Method of one of claims 13 to ~~18~~17, comprising determining an estimated adaptive codebook distortion ~~(116)~~ which an adaptive codebook used in the second coding algorithm would introduce when using the adaptive codebook to encode the portion of the audio signal, and estimating the second quality measure based on an energy of a portion of a weighted version of the audio signal and the estimated adaptive codebook distortion, and

comprising, for each of a plurality of sub-portions of the portion of the audio signal, approximating ~~(112)~~ the adaptive codebook based on a version of the sub-portion of the weighted audio signal shifted to the past by a pitch-lag determined in a pre-

processing stage, estimating ~~(114)~~ an adaptive codebook gain such that an error between the sub-portion of the portion of the weighted audio signal and the approximated adaptive codebook is minimized, and determining ~~(116)~~ the estimated adaptive codebook distortion based on the energy of an error between the sub-portion of the portion of the weighted audio signal and the approximated adaptive codebook scaled by the adaptive codebook gain.

19. Method of claim 18, comprising reducing ~~(118)~~ the estimated adaptive codebook distortion determined for each sub-portion of the portion of the audio signal by a constant factor.

20. Method of claim 18 or 19, wherein the second quality measure is a segmental SNR of the portion of the weighted audio signal, and comprising estimating the segmental SNR by calculating an estimated SNR associated with each sub-portion based on the energy of the corresponding sub-portion of the weighted audio signal and the estimated adaptive codebook distortion and by calculating an average of the SNRs associated with the sub-portions to obtain the estimated segmental SNR for the portion of the weighted audio signal.

21. Method of claim 18, comprising approximating the adaptive codebook based on a version of the portion of the weighted audio signal shifted to the past by a pitch-lag determined in a pre-processing stage, estimating an adaptive codebook gain such that an error between the portion of the weighted audio signal and the approximated adaptive codebook is minimized, and determining the estimated adaptive codebook distortion based on the energy of an error between the portion of the weighted audio signal and the approximated adaptive codebook scaled by the adaptive codebook gain.

22. Method of one of claims 13 to 21~~22~~, comprising utilizing a hysteresis in comparing the estimated quality measures.

23. Computer program having a program code for performing, when running on a computer, the method of one of claims 13 to 22.