frequencies with quantized values unequal to zero, leaving dead-zones that are not further improved.

On the other hand, an underestimation of the bit-consumption inevitably leads to partial loss of spectral coefficients, usually the highest frequencies. In state of the art encoders this effect is mitigated by applying noise substitution at the decoder, which is based on the assumption that high frequency content is usually noisy.

In this setup it is evident, that it is desirable to encode as much of the signal as possible in the first encoding step, which uses entropy coding and is therefore more efficient than the residual coding step. Therefore, one would like to select the global gain with a bit estimate as close to the available bit-budget as possible. While the power spectrum based estimator works well for most audio content, it can cause problems for highly tonal signals, where the first stage estimation is mainly based on irrelevant side-lobes of the frequency decomposition of the filter-bank while important components are lost due to underestimation of the bit-consumption.

It is the object of the present invention to provide an improved concept for audio encoding or decoding, that, nevertheless, is efficient and obtains a good audio quality.

This object is achieved by an audio encoder of claim 1, and a method of encoding audio input data of claim 33, or a computer program of claim 35.

The present invention is based on the finding that, in order to enhance the efficiency particularly with respect to the bitrate on the one hand and the audio quality on the other hand, a signal-dependent change with respect to the typical situation that is given by psychoacoustic considerations is necessary. Typical psychoacoustic models or psychoacoustic considerations result in a good audio quality at a low bitrate for all signal classes in average, i.e., for all audio signal frames irrespective of their signal characteristic, when an average result is contemplated. However, it has been found that for certain signal classes or for signals having certain signal characteristics such as quite tonal signals, the straightforward psychoacoustic model or the straight forward psychoacoustic control of the encoder only results in sub-optimum outcomes with respect to audio quality (when the bitrate is kept constant), or with respect to bitrate (when the audio quality is kept constant).

Therefore, in order to address this shortcoming of typical psychoacoustic considerations, the present invention provides, in the context of an audio encoder with a preprocessor for preprocessing the audio input data to obtain audio data to be encoded, and a coder processor for coding the audio data to be coded, a controller for controlling the coder processor in such a way that, depending on a certain signal characteristic of a frame, a number of audio data items of the audio data to be coded by the coder processor is reduced compared to typical straightforward results obtained by state of the art psychoacoustic considerations. Furthermore, this reduction of the number of audio data items is done in a signal-dependent way so that, for a frame with a certain first signal characteristic, the number is stronger reduced than for another frame with another signal characteristic that differs from the signal characteristic from the first frame. This reduction in the number of audio data items can be considered to be a reduction in the absolute number or a reduction in the relative number, although this is not decisive. It is, however, a feature that the information units that are “saved” by the intentional reduction of the number of audio data items are not simply lost, but are used for more precisely coding the remaining number of data items, i.e., the data items that have not been eliminated by the intentional reduction of the number of audio data items.

In accordance with a first aspect of the invention, there is provided an audio encoder for encoding audio input data, comprising: a preprocessor for preprocessing the audio input data to obtain audio data to be coded; a coder processor for coding the audio data to be coded; and a controller for controlling the coder processor so that, depending on a first signal characteristic of a first frame of the audio data to be coded, a number of audio data items of the audio data to be coded by the coder processor for the first frame is reduced compared to a second signal characteristic of a second frame, and a first number of information units used for coding the reduced number of audio data items for the first frame is stronger enhanced compared to a second number of information units for the second frame.

In accordance with a second aspect of the invention, there is provided a method of encoding audio input data, comprising: preprocessing the audio input data to obtain audio data to be coded; coding the audio data to be coded; and controlling the coding so that, depending on a first signal characteristic of a first frame of the audio data to be coded, a number of audio data items of the audio data to be coded for the first frame is reduced compared to a second signal characteristic of a second frame, and a first number of information units used for

coding the reduced number of audio data items for the first frame is stronger enhanced compared to a second number of information units for the second frame.

In accordance with the invention, the controller for controlling the coder processor operates in such a way that, depending on the first signal characteristic of a first frame of the audio data to be coded, a number of audio data items of the audio data to be coded by the coder processor for the first frame is reduced compared to a second signal characteristic of a second frame, and, at the same time, a first number of information units used for coding the reduced number of audio data items for the first frame is stronger enhanced compared to a second number of information units for the second frame.

In a preferred embodiment, the reduction is done in such a way that, for more tonal signal frames, a stronger reduction is performed and, at the same time, the number of bits for the individual lines is stronger enhanced compared to a frame that is less tonal, i.e., that is more noisy. Here, the number is not reduced to such a high degree and, correspondingly, the number of information units used for encoding the less tonal audio data items is not increased so much.

The present invention provides a framework where, in a signal dependent way, typically provided psychoacoustic considerations are more or less violated. On the other hand

Claims

1. Audio encoder for encoding audio input data, comprising:

a preprocessor for preprocessing the audio input data to obtain audio data to be coded;

a coder processor for coding the audio data to be coded; and

a controller for controlling the coder processor so that, depending on a first signal characteristic of a first frame of the audio data to be coded, a number of audio data items of the audio data to be coded by the coder processor for the first frame is reduced compared to a second signal characteristic of a second frame, and a first number of information units used for coding the reduced number of audio data items for the first frame is stronger enhanced compared to a second number of information units for the second frame.

2. Audio encoder of claim 1,

wherein the coder processor comprises an initial coding stage and a refinement coding stage,

wherein the controller is configured to reduce the number of audio data items encoded by the initial coding stage for the first frame,

wherein the initial coding stage is configured to code the reduced number of audio data items for the first frame using a first frame initial number of information units, and

wherein the refinement coding stage is configured to use a first frame remaining number of information units for a refinement coding for the reduced number of audio data items for the first frame, wherein the first frame initial number of information units added to the first frame remaining number of information units result in a predetermined number of information units for the first frame.

3. Audio encoder of claim 2,

wherein the controller is configured to reduce the number of audio data items encoded by the initial coding stage for the second frame to a higher number of audio data items compared to the first frame,

wherein the initial coding stage is configured to code the reduced number of audio data items for the second frame using a second frame initial number of information units, the second frame initial number of information units being higher than the first frame initial number of information units, and

wherein the refinement coding stage is configured to use a second frame remaining number of information units for a refinement coding for the reduced number of audio data items for the second frame, wherein the second frame initial number of information units added to the second frame remaining number of information units result in the predetermined number of information units for the first frame.

4. Audio encoder of claim 1,

wherein the coder processor comprises an initial coding stage and a refinement coding stage,

wherein the initial coding stage is configured to code the reduced number of audio data items for the first frame using a first frame initial number of information units,

wherein the refinement coding stage is configured to use a first frame remaining number of information units for a refinement coding for the reduced number of audio data items for the first frame, wherein the first frame initial number of information units added to the first frame remaining number of information units result in a predetermined number of information units for the first frame, and

wherein the controller is configured to control the coder processor so that the refinement coding stage performs a refinement coding of at least one of the reduced number of audio data items of the first frame using at least two information units, or so that the refinement coding stage performs a refinement coding of more than 50 percent of the reduced number of audio data items using at least two information units for each audio data item, or

wherein the controller is configured to control the coder processor so that the refinement coding stage performs a refinement coding of all audio data items of the second frame using less than two information units, or so that the refinement coding stage performs a refinement coding of less than 50 percent of the reduced number of audio data items using at least two information units for each audio data item.

5. Audio encoder of claim 1,

wherein the coder processor comprises an initial coding stage and a refinement coding stage,

wherein the initial coding stage is configured to code the reduced number of audio data items for the first frame using a first frame initial number of information units,

wherein the refinement coding stage is configured to use a first frame remaining number of information units for a refinement coding for the reduced number of audio data items for the first frame,

wherein the refinement coding stage is configured to iteratively assign the first frame remaining number of information units to the reduced number of audio data items in at least two sequentially performed iterations, to calculate values of the assigned information units for the at least two sequentially performed iterations and to introduce the calculated values of the information units for the at least two sequentially performed iterations into an encoded output frame in a predetermined order.

6. Audio encoder of claim 5, wherein the refinement coding stage is configured to sequentially calculate an information unit for each audio data item of the reduced number of audio data items for the first frame in an order from a low frequency information for the audio data item to a high frequency information for the audio data item in a first iteration,

wherein the refinement coding stage is configured to sequentially calculate an information unit for each audio data item of the reduced number of audio data items

for the first frame in an order from a low frequency information for the audio data item to a high frequency information for the audio data item in a second iteration, and

wherein the refinement coding stage is configured to check, whether a number of already assigned information units is lower than a predetermined number of information units for the first frame less than the first frame initial number of information units and to stop the second iteration in case of a negative check result, or in case of a positive check result, to perform a number of further iterations, until a negative check result is obtained, the number of further iterations being at least one, or

wherein the refinement coding stage is configured to count a number of non-zero audio items, and to determine the number of iterations from the number of non-zero audio items and a predetermined number of information units for the first frame less than the first frame initial number of information units.

7. Audio encoder of claim 1,

wherein the coder processor comprises an initial coding stage and a refinement coding stage,

wherein the initial coding stage is configured to code a number of most significant information units for each audio data item of the reduced number of audio data items for the first frame using a first frame initial number of information units, the number being greater than one, and

wherein the refinement coding stage is configured to use a first frame remaining number of information units for encoding a number of least significant information units for each audio data item of the reduced number of audio data items for the first frame, the number being greater than one for at least one audio data item of the reduced number of audio data items for the first frame.

8. Audio encoder of any one of the preceding claims,

wherein the first signal characteristic is a first tonality value, wherein the second signal characteristic is a second tonality value, and wherein the first tonality value indicates a higher tonality than the second tonality value, and

wherein the controller is configured to reduce the number of audio data items for the first frame to a first number being smaller than the number of audio data items for the second frame, and to increase an average number of information units used for coding each audio data item of the reduced number of audio data items of the first frame to be greater than an average number of information units used for coding each audio data item of the reduced number of audio data items of the second frame.

9. Audio encoder of claim 1, wherein the coder processor comprises:

a variable quantizer for quantizing the audio data of the first frame to obtain quantized audio data for the first frame and for quantizing the audio data of the second frame to obtain quantized audio data for the second frame;

an initial coding stage for coding the quantized audio data of the first frame or the second frame;

a refinement coding stage for encoding residual data of the first frame and the second frame;

wherein the controller is configured for analyzing the audio data of the first frame to determine a first control value for the variable quantizer for the first frame and for analyzing the audio data of the second frame to determine a second control value for the variable quantizer for the second frame, the second control value being different from the first control value, and

wherein the controller is configured to perform a manipulation of the audio data of the first frame or the second frame or of amplitude-related values derived from the audio data of the first frame or the second frame depending on the audio data for determining the first control value or the second control value, and wherein the variable quantizer is configured to quantize the audio data of the first frame or the second frame without the manipulation.

10. Audio encoder of claim 1, wherein the coder processor comprises:

a variable quantizer for quantizing the audio data of the first frame to obtain quantized audio data for the first frame and for quantizing the audio data of the second frame to obtain quantized audio data for the second frame;

an initial coding stage for coding the quantized audio data of the first frame or the second frame;

a refinement coding stage for encoding residual data of the first frame and the second frame;

wherein the controller is configured for analyzing the audio data of the first frame to determine a first control value for the variable quantizer, for the initial coding stage or for an audio data item reducer for the first frame and for analyzing the audio data of the second frame to determine a second control value for the variable quantizer, for the initial coding stage or for an audio data item reducer for the second frame, the second control value being different from the first control value, and

wherein the controller is configured to determine a first tonality characteristic as the first signal characteristic to determine the first control value, and a second tonality characteristic as the second signal characteristic to determine the second control value so that a bit-budget for the refinement coding stage is increased in case of a first tonality characteristic compared to the bit-budget for the refinement coding stage in case of a second tonality characteristic, wherein the first tonality characteristic indicates a greater tonality than the second tonality characteristic.

11. Audio encoder of claim 9 or 10, wherein the initial coding stage is an entropy coding stage for entropy coding, or wherein the refinement coding stage is a residual coding stage or a binary coding stage for encoding residual data of the first frame and the second frame.

12. Audio encoder of any one of claims 9 to 11,

wherein the controller is configured to determine the first control value or the second control value so that a first budget of information units for the initial coding stage is lower than or equal to a predefined value, and wherein the controller is configured to derive a second budget of information units for the refinement coding stage using the first budget of information units and the maximum number of information units for the first or second frame or the predefined value.

13. Audio encoder of any one of claims 9 to 12, wherein the controller is configured to calculate the amplitude-related values as a plurality of power values derived from one or more audio values of the audio data and to manipulate the power values using an addition of an identical manipulation value to all power values of the plurality of power values, or

wherein the controller is configured

to randomly add or subtract an identical manipulation value to or from all audio values of a plurality of audio values included in the frame, or,

to add or subtract values obtained by the same magnitude of the manipulation value, or

to add or subtract values obtained by a subtraction of slightly different terms from the same magnitude

to add or subtract values obtained as samples from a normalized probability distribution scaled using a calculated complex or real magnitude of the manipulation value, or

wherein the controller is configured to calculate the amplitude-related values using an exponentiation of the audio data of the first or second frame or of downsampled audio data of the first or second frame with an exponent value, the exponent value being greater than 1.

14. Audio encoder of claim 9, wherein the controller is configured to calculate a manipulation value for the manipulation using a maximum value of the plurality of audio data or of the amplitude-related values or using a maximum value of a

plurality of downsampled audio data or a plurality of downsampled amplitude-related values for the first or second frame.

15. Audio encoder of claim 9, wherein the controller is configured to calculate a manipulation value for the manipulation additionally using a signal independent weighting value, the signal independent weighting value depending on at least one of a bit-rate for the first or second frame, a frame duration, and a sampling frequency.

16. Audio encoder of any claim 9, wherein the controller is configured to calculate a manipulation value for the manipulation using a signal dependent weighting value derived from at least one of a first sum of magnitudes of the audio data or downsampled audio data within the frame, a second sum of magnitudes of the audio data or the downsampled audio data within the frame multiplied by an index associated with each magnitude, and a quotient of the second sum and the first sum.

17. Audio encoder of claim 9,

wherein the controller is configured to calculate a manipulation value for the manipulation based on the following equation:

$$N(X_f) = \max_{\mathrm{k}} \left| X_f(k) \right| * 2^{-regBits-lowBits}$$

wherein k is a frequency index, wherein X_f (k) is an audio data value for the frequency index k before quantization, wherein max is the maximum function, wherein *regBits* is a first signal independent weighting value, and wherein *lowBits* is a second signal dependent weighting value.

18. Audio encoder of any one of the preceding claims, wherein the preprocessor further comprises:

a time-frequency converter for converting time domain audio data into spectral values of the frame; and

a spectral processor for calculating modified spectral values having a spectral envelope being flatter than a spectral envelope of the spectral values, wherein the modified spectral values represent the audio data of the first or the second frame to be encoded by the coder processor.

19. Audio encoder of claim 18, wherein the spectral processor is configured to perform at least one of a temporal noise shaping operation, a spectral noise shaping operation, and a spectral whitening operation.

20. Audio encoder of claim 9, wherein the controller is configured to calculate the first control value or the second control value using a plurality of energy values as the amplitude-related values for the frame, wherein each energy value of the plurality of energy values is derived from a power value as an amplitude-related value of the amplitude-related values for the frame and a signal-dependent manipulation value for the manipulation.

21. Audio encoder of claim 20, wherein the controller is configured

to calculate a required bit estimate of each energy value of the plurality of energy values depending on the energy value and a candidate value for the first control value or the second control value,

to accumulate the required bit estimates for the energy values of the plurality of energy values and the candidate value for the first control value or the second control value,

to check, whether an accumulated bit estimate for the candidate value for the first control value or the second control value fulfills an allowed bit consumption criterion, and

to modify the candidate value for the control value in case the allowed bit consumption criterion is not fulfilled and to repeat the calculation of the required bit estimate, the accumulation of the bit estimate and the checking until a fulfillment of the allowed bit consumption criterion for a modified candidate value for the first control value or the second control value is found.

22. Audio encoder of claim 20 or 21,

wherein the controller is configured to calculate the plurality of energy values based on the following equation:

$$E(k) = 10\log_{10}\left(PX_{lp}(k) + N\left(X_f\right) + 2^{-31}\right),$$

wherein *E(k)* is an energy value of the plurality of energy values for an index *k,* wherein $PX_{lp}(k)$ is a power value for an index *k* as the amplitude related value, and wherein $N(X_f)$ is the signal dependent manipulation value.

23. Audio encoder of any one of the claims 9 to 22, wherein the controller is configured to calculate the first control value or the second control value based on an estimation of accumulated information units required for each manipulated audio data value or manipulated amplitude-related value.

24. Audio encoder of claim 9 or 10,

wherein the controller is configured to manipulate in such a way that due to manipulation, a bit-budget for the initial coding stage is increased or a bit-budget for the refinement coding stage is decreased.

25. Audio encoder of claim 9 or 10,

wherein the controller is configured to manipulate in such a way that the manipulation results in a higher bit-budget of the refinement coding stage for a signal with a first tonality compared to a signal with a second tonality, wherein the second tonality is lower than the first tonality.

26. Audio encoder of claim 9 or 10,

wherein the controller is configured to manipulate in such a way that an energy of the audio data, from which a bit-budget for the initial coding stage is calculated, is increased with respect to the energy of the audio data to be quantized by the variable quantizer.

27. Audio encoder of any one of the claims 1 to 8, wherein the coder processor comprises a variable quantizer for quantizing the audio data of the first frame to obtain quantized audio data for the first frame and for quantizing the audio data of the second frame to obtain quantized audio data for the second frame,

wherein the controller is configured to calculate a global gain for the first or the second frame, and

wherein the variable quantizer comprises: a weighter for weighting with the global gain; and a quantizer core having a fixed quantization step size.

28. Audio encoder of claim 9,
wherein the refinement coding stage is configured for calculating refinement bits for quantized audio values in a plurality of iterations, wherein, in each iteration, a refinement bit indicates a different amount, or

wherein a refinement bit in a lower iteration indicates a higher amount than a refinement bit in a higher iteration, or

wherein the amount is a fractional amount being a fraction of a quantizer step size indicated by the first control value or the second control value.

29. Audio encoder of claim 6wherein the refinement coding stage is configured

to perform an iterative processing having at least two iterations,

to check, whether a quantized audio value or the quantized audio value together with a potential first amount associated with a refinement bit for the quantized audio value in a first iteration, added to or subtracted from a second amount for the second iteration when weighted by a global gain is greater than or lower than a non-quantized audio value, and

to set a refinement bit for the second iteration depending on a result of the check.

30. Audio encoder of any one of the claims 1 to 8, wherein the coder processor comprises a variable quantizer and a refinement coding stage, wherein the refinement coding stage is configured to calculate a refinement bit only for audio values that are not quantized to zero by the variable quantizer.

31. Audio encoder of claim 1, wherein the coder processor comprises an initial coding stage,

wherein the controller is configured to reduce an impact of a manipulation for the audio data having a center of mass at a lower frequency, and

wherein the initial coding stage is configured to remove high frequency spectral values from the audio data in case it is determined that a bit-budget for the first or the second frame does not suffice for encoding quantized audio data of the frame.

32. Audio encoder of any one of the preceding claims,

wherein the controller is configured to perform a bi-section search for each frame individually using manipulated spectral energy values for the first frame or the second frame as manipulated amplitude-related values for the first or the second frame.

33. Method of encoding audio input data, comprising:

preprocessing the audio input data to obtain audio data to be coded;

coding the audio data to be coded; and

controlling the coding so that, depending on a first signal characteristic of a first frame of the audio data to be coded, a number of audio data items of the audio data to be coded for the first frame is reduced compared to a second signal characteristic of a second frame, and a first number of information units used for coding the reduced number of audio data items for the first frame is stronger enhanced compared to a second number of information units for the second frame.

34. Method of claim 33, wherein the coding comprises:

variably quantizing audio data of a frame to obtain quantized audio data;

entropy coding the quantized audio data of the frame; and

encoding residual data of the frame;

wherein the controlling comprises determining a control value for the variably quantizing, the determining comprising: analyzing the audio data of the first or the second frame; and performing a manipulation of the audio data of the first or the second frame or amplitude-related values derived from the audio data of the first or the second frame depending on the audio data for determining the control value, wherein the variably quantizing quantizes the audio data of the frame without the manipulation, or

wherein the controlling comprises determining a first or second tonality characteristic of the audio data and determining the control value so that a bit-budget for encoding the residual data is increased in case of the first tonality characteristic compared to the bit-budget for encoding the residual data in case of the second tonality characteristic, wherein the first tonality characteristic indicates a greater tonality than the second tonality characteristic.

35. Computer program for performing, when running on a computer or a processor, the method of claim 33 or claim 34.