

## Information Signal Encoding

### Description

5 The present invention relates to information signal encoding, such as audio or video encoding.

The usage of digital audio encoding in new communication networks as well as in professional audio productions for  
10 bi-directional real time communication requires a very inexpensive algorithmic encoding as well as a very short encoding delay. A typical scenario where the application of digital audio encoding becomes critical in the sense of the delay time exists when direct, i.e. unencoded, and  
15 transmitted, i.e. encoded and decoded signals are used simultaneously. Examples therefore are live productions using cordless microphones and simultaneous (in-ear) monitoring or "scattered" productions where artists play simultaneously in different studios. The tolerable overall  
20 delay time period in these applications is less than 10 ms. If, for example, asymmetrical participant lines are used for communication, the bit rate is an additional limiting factor.

25 The algorithmic delay of standard audio encoders, such as MPEG-1 3 (MP3), MPEG-2 AAC and MPEG-2/4 low delay ranges from 20 ms to several 100 ms, wherein reference is made, for example, to the article M. Lutzky, G. Schuller, M. Gayer; U. Kraemer, S. Wabnik: "A guideline to audio codec  
30 delay", presented at the 116<sup>th</sup> AES Convention, Berlin, May 2004. Voice encoders operate at lower bit rates and with less algorithmic delay, but provide merely a limited audio quality.

35 The above outlined gap between the standard audio encoders on the one hand and the voice encoders on the other hand is, for example, closed by a type of encoding scheme described in the article B. Edler, C. Faller and G.

Schuller, "Perceptual Audio Coding Using a Time-Varying Linear Pre- and Postfilter", presented at 109<sup>th</sup> AES Convention, Los Angeles, September 2000, according to which the signal to be encoded is filtered with the inverse of the marking threshold on the encoder side and is subsequently quantized to perform irrelevance reduction, and the quantized signal is supplied to entropy encoding for performing redundancy reduction separate from the irrelevance reduction, while the quantized prefiltered signal is reconstructed on the decoder side and filtered in a postfilter with the marking threshold as transmission function. Such an encoding scheme, referred to as ULD encoding scheme below, results in a perceptual quality that can be compared to standard audio encoders, such as MP3, for bit rates of approximately 80 kBit/s per channel and higher. An encoder of this type is, for example, also described in WO 2005/078703 A1.

Particularly, the ULD encoders described there use psychoacoustically controlled linear filters for forming the quantizing noise. Due to their structure, the quantizing noise is always on the given threshold, even when no signal is in a given frequency domain. The noise remains inaudible, as long as it corresponds to the psychoacoustic masking threshold. For obtaining a bit rate that is even smaller than the bit rate as predetermined by this threshold, the quantizing noise has to be increased, which makes the noise audible. Particularly, the noise becomes audible in domains without signal portions. Examples therefore are very low and very high audio frequencies. Normally, there are only very low signal portions in these domains, while the masking threshold is high. If the masking threshold is increased uniformly across the whole frequency domain, the quantizing noise is at the increased threshold, even when there is no signal, so that the quantizing noise becomes audible as a signal that sounds spurious. Subband-based encoders do not have

this problem, since the same simply quantize subbands having smaller signals than the threshold to zero.

5 The above-mentioned problem that occurs when the allowed bit rate falls below the minimum bit rate, which causes no spurious quantizing noise and which is determined by the masking threshold, is not the only one. Further, the ULD encoders described in the above references suffer from a complex procedure for obtaining a constant data rate,  
10 particularly since an iteration loop is used, which has to be passed in order to determine, per sampling block, an amplification factor value adjusting a dequantizing step size.

15 It is the object of the present invention to provide an information encoding scheme that makes it possible to allow the short delay time typical for ULD encoder types at a low bit rate and yet a high encoding quality.

20 This object is achieved by apparatuses according to claim 1 or 24, methods according to claim 44 or 45 as well as an encoder according to claim 47 and a decoder according to claim 48.

25 The central idea of the present invention is the finding that extremely coarse quantization exceeding the measure determined by the masking threshold is made possible, without or only very little quality losses, by not directly quantizing the prefiltered signal but a prediction error  
30 obtained by forward-adaptive prediction of the prefiltered is. Due to the forward adaptivity, the quantizing error has no negative effect on the prediction coefficient.

According to a further embodiment, the prefiltered signal  
35 is even quantized in a nonlinear manner or even clipped, i.e. quantized via a quantizing function, which maps the unquantized values of the prediction error on quantizing indices of quantizing stages, and whose course is steeper

below a threshold than above a threshold. Thereby, the noise PSD increased in relation to the masking threshold due to the low available bit rate adjusts to the signal PSD, so that the violation of the masking threshold  
5 does not occur at spectral parts without signal portion, which further improves the listening quality or maintains the listening quality, respectively, despite a decreasing available bit rate.

10 According to a further embodiment of the present invention, quantization is even quantized or limited, respectively, by clipping, namely by quantizing to a limited and fixed number of quantizing levels or stages, respectively. By prediction of the prefiltered signal via forward-adaptive  
15 prediction, the coarse quantization has no negative effect on the prediction coefficients themselves. By quantizing to a fixed number of quantizing levels, prevention of iteration for obtaining a constant bit rate is inherently enabled.

20 According to a further embodiment of the present invention, a quantizing step size or stage height, respectively, between the fixed number of quantizing levels is determined in a backward-adaptive manner from previous quantizing  
25 level indices obtained by quantization, so that, on the one hand, despite a very low number of quantizing levels, a better or at least best possible quantization of the prediction error or residual signal, respectively, can be obtained, without having to provide further side  
30 information to the decoder side. On the other hand, it is possible to ensure that transmission errors during transmission of the quantized residual signal to the decoder side only have a short-time effect on the decoder side with appropriate configuration of the backward-  
35 adaptive step size adjustment.

Preferred embodiments of the invention will be discussed below with reference to the accompanying drawings. They show:

- 5    Fig. 1      a block diagram of an encoder according to an embodiment of the present invention;

10           Figs. 2a/b graphs showing exemplarily the course of the noise spectrum in relation to the masking threshold and signal power spectrum density for the case of the encoder according to claim 1 (graph a) or for a comparative case of an encoder with backward-adaptive prediction of the prefiltered signal and iterative and masking threshold block-wise quantizing step size adjustment (graph b), respectively;

20           Figs. 3a/3b and 3c graphs showing exemplarily the signal power spectrum density in relation to the noise or error power spectrum density, respectively, for different clip extensions or different numbers of quantizing levels, respectively, for the case that, like in the encoder of Fig. 1, forward-adaptive prediction of the prefiltered signal but still an iterative quantizing step size adjustment is performed;

- 30           Fig. 4      a block diagram of a structure of the coefficient encoder in the encoder of Fig. 1 according to an embodiment of the present invention;

- 35           Fig. 5      a block diagram of a decoder for decoding an information signal encoded by the encoder of Fig. 1 according to an embodiment of the present invention;

- Fig. 6      a block diagram of a structure of the coefficient encoders in the encoder of Fig. 1 or the decoder

of Fig. 5 according to an embodiment of the present invention;

Fig. 7 a graph for illustrating listening test results;  
5 and

Figs. 8a to 8c graphs of exemplary quantizing functions  
that can be used in the quantizing and  
quantizing/clip means, respectively, in Figs. 1,  
10 4, 5 and 6.

Before embodiments of the present invention will be  
discussed in more detail with reference to the drawings,  
first, for a better understanding of the advantages and  
15 principles of these embodiments, a possible implementation  
of an ULD-type encoding scheme will be discussed as  
comparative example, based on which the essential  
advantages and considerations underlying the subsequent  
embodiments, which have finally led to these embodiments,  
20 can be illustrated more clearly.

As has already been described in the introduction of the  
description, there is a need for an ULD version for lower  
bit rates of, for example, 64 k Bit/s, with comparable  
25 perceptual quality, as well as simpler scheme for obtaining  
a constant bit rate, particularly for intended lower bit  
rates. Additionally, it would be advantageous when the  
recovery time after a transmission error would remain low  
or at a minimum.

30 For redundancy reduction of the psychoacoustically  
preprocessed signal, the comparison ULD encoder uses a  
sample-wise backward-adaptive closed-loop prediction. This  
means that the calculation of prediction coefficients in  
35 encoder and decoder is based merely on past or already  
quantized and reconstructed signal samples. For obtaining  
an adaption to the signal or the prefiltered signal,  
respectively, a new set of predictor coefficients is

calculated again for every sample. This results in the advantage that long predictors or prediction value determination formulas, i.e. particularly predictors having a high number of predictor coefficients can be used, since  
5 there is no requirement to transmit the predictor coefficients from encoder to decoder side. On the other hand, this means that the quantized prediction error has to be transmitted to the decoder without accuracy losses, for obtaining prediction coefficients that are identical to  
10 those underlying the encoding process. Otherwise, the predicted or predicated values, respectively, in the encoder and decoder would not be identical to each other, which would cause an instable encoding process. Rather, in the comparison ULD encoder, periodical reset of the  
15 predictor both on encoder and decoder side is required to allow selective access to the encoded bit stream as well as to stop a propagation of transmission errors. However, the periodic resets cause bit rate peaks, which presents no problem for a channel with variable bit rate, but for  
20 channels with fixed bit rate where the bit rate peaks limit the lower limit of a constant bit rate adjustment.

As will result from the subsequent more detailed description of the ULD comparison encoding scheme with the  
25 embodiments of the present invention, these embodiments differ from the comparison encoding scheme by using a block-wise forward-adaptive prediction with a backward-adaptive quantizing step size adjustment instead of a sample-wise backward-adaptive prediction. On the one hand,  
30 this has the disadvantage that the predictors should be shorter in order to limit the amount of required side information for transmitting the required prediction coefficients towards the encoder side, which again might result in reduced encoder efficiency, but, on the other  
35 hand, this has the advantage that the procedure of the subsequent embodiments still functions effectively for higher quantizing errors, which are a result of reduced bit

rates, so that the predictor on the decoder side can be used for quantizing noise shaping.

As will also result from the subsequent comparison, compared to the comparison ULD encoder, the bit rate is limited by limiting the range of values of the prediction remainder prior to transmission. This results in noise shaping modified compared to the comparison ULD encoding scheme, and also leads to different and less spurious listening artifacts. Further, a constant bit rate is generated without using iterative loops. Further, "reset" is inherently included for every sample block as result of the block-wise forward adaption. Additionally, in the embodiments described below, an encoding scheme is used for prefilter coefficients and forward prediction coefficients, which uses difference encoding with backward-adaptive quantizing step size control for an LSF (line spectral frequency) representation of the coefficients. The scheme provides block-wise access to the coefficients, generates a constant side information bit rate and is, above that, robust against transmission errors, as will be described below.

In the following, the comparison ULD encoder and decoder structure will be described in more detail, followed by the description of embodiments of the present invention and the illustration of its advantages in the transmission from higher constant bit rates to lower bit rates.

In the comparison ULD encoding scheme, the input signal of the encoder is analyzed on the encoder side by a perceptual model or listening model, respectively, for obtaining information about the perceptually irrelevant portions of the signal. This information is used to control a prefilter via time-varying filter coefficients. Thereby, the prefilter normalizes the input signal with regard to its masking threshold. The filter coefficients are calculated

once for every block of 128 samples each, quantized and transmitted to the encoder side as side information.

After multiplication of the prefiltered signal with an  
5 amplification factor by subtracting the backward-adaptive  
predicted signal, the prediction error is quantized by a  
uniform quantizer, i.e. a quantizer with uniform step size.  
As already mentioned above, the predicted signal is  
obtained via sample-wise backward-adaptive closed-loop  
10 prediction,. Accordingly, no transmission of prediction  
coefficients to the decoder is required. Subsequently, the  
quantized prediction residual signal is entropy encoded.  
For obtaining a constant bit rate, a loop is provided,  
which repeats the steps of multiplication, prediction,  
15 quantizing and entropy-encoding several times for every  
block of prefiltered samples. After iteration, the highest  
amplification factor of a set of predetermined  
amplification values is determined, which still fulfills  
the constant bit rate condition. This amplification value  
20 is transmitted to the decoder. If, however, an  
amplification value smaller than one is determined, the  
quantizing noise is perceptible after decoding, i.e. its  
spectrum is shaped similar to the masking threshold, but  
its overall power is higher than predetermined by the  
25 prediction model. For portions of the input signal  
spectrum, the quantizing noise could even get higher than  
the input signal spectrum itself, which again generates  
audible artifacts in portions of the spectrum, where  
otherwise no audible signal would be present, due to the  
30 usage of a predictive encoder. The effects caused by  
quantizing noise represent a limiting factor when lower  
constant bit rates are of interest.

Continuing with the description of the comparison ULD  
35 scheme, the prefilter coefficients are merely transmitted  
as intraframe LSF differences, and also only as soon as the  
same exceed a certain limit. For avoiding transmission  
error propagation for an unlimited period, the system is

reset from time to time. Additional techniques can be used for minimizing a decrease in perception of the decoded signal in the case of transmission errors. The transmission scheme generates a variable side information bit rate, which is leveled in the above-described loop by adjusting the above-mentioned amplification factor accordingly.

The entropy encoding of the quantized prediction residual signal in the case of the comparison ULD encoder comprises methods, such as a Golomb, Huffman, or arithmetic encoding method. The entropy encoding has to be reset from time to time and generates inherently a variable bit rate, which is again leveled by the above-mentioned loop.

In the case of the comparison ULD encoding scheme, the quantized prediction residual signal in the decoder is obtained from entropy encoding, whereupon the prediction remainder and the predicted signal are added, the sum is multiplied with the inverse of the transmitted amplification factor, and therefrom, the reconstructed output signal is generated via the postfilter having a frequency response inverse to the one of the prefilter, wherein the postfilter uses the transmitted prefilter coefficients.

25

A comparison ULD encoder of the just described type obtains, for example, an overall encoder/decoder delay of 5.33 to 8 ms at sample frequencies of 32 kHz to 48 kHz. Without (spurious loop) iterations, the same generates bit rates in the range of 80 to 96 kBit/s. As described above, at lower constant bit rates, the listening quality is decreased in this encoder, due to the uniform increase of the noise spectrum. Additionally, due to the iterations, the effort for obtaining a uniform bit rate is high. The embodiments described below overcome or minimize these disadvantages. At a constant transmission data rate, the encoding scheme of the embodiments described below causes altered noise shaping of the quantizing error and requires

no iteration. More precisely, in the above-discussed comparison ULD encoding scheme, in the case of constant transmission data rate in an iterative process, a multiplier is determined, with the help of which the  
5 signal coming from the prefilter is multiplied prior to quantizing, wherein the quantizing noise is spectrally white, which causes a quantizing noise in the decoder which is shaped like the listening threshold, but which lies slightly below or slightly above the listening threshold,  
10 depending on the selected multiplier, which can, as described above, also be interpreted as a shift of the determined listening threshold. In connection therewith, quantizing noise results after decoding, whose power in the individual frequency domains can even exceed the power of  
15 the input signal in the respective frequency domain. The resulting encoding artifacts are clearly audible. The embodiments described below shape the quantizing noise such that its spectral power density is no longer spectrally white. The coarse quantizing/limiting or clipping,  
20 respectively, of the prefilter signal rather shapes the resulting quantizing noise similar to the spectral power density of the prefilter signal. Thereby, the quantizing noise in the decoder is shaped such that it remains below the spectral power density of the input signal. This can be  
25 interpreted as deformation of the determined listening threshold. The resulting encoding artifacts are less spurious than in the comparison ULD encoding scheme. Further, the subsequent embodiments require no iteration process, which reduces complexity.

30

Since by describing the comparison ULD encoding scheme above, a sufficient base has been provided for turning the attention to the underlying advantages and considerations of the following embodiments for the description of these  
35 embodiments, first, the structure of an encoder according to an embodiment of the present invention will be described below.

The encoder of Fig. 1, generally indicated by 10, comprises an input 12 for the information signal to be encoded, as well as an output 14 for the encoded information signal, wherein it is exemplarily assumed below that this is an audio signal, and exemplarily particularly an already sampled audio signal, although sampling within the encoder subsequent to the input 12 would also be possible. Samples of the incoming output signal are indicated by  $x(n)$  in Fig. 1.

10

As shown in Fig. 1, the encoder 10 can be divided into a masking threshold determination means 16, a prefilter means 18, a forward-predictive prediction means 20 and a quantizing/clip means 22 as well as bit stream generation means 24. The masking threshold determination means 16 operates according to a perceptual model or listening model, respectively, for determining a representation of the masking or listening threshold, respectively, of the audio signal incoming at the input 12 by using the perceptual model, which indicates a portion of the audio signal that is irrelevant with regard to the perceptibility or audibility, respectively, or represents a spectral threshold for the frequency at which which spectral energy remains inaudible due to psychoacoustic covering effects or is not perceived by humans, respectively. As will be described below, the determining means 16 determines the masking threshold in a block-wise manner, i.e. the same determines a masking threshold per block of subsequent blocks of samples of the audio signal. Other procedures would also be possible. The representation of the masking threshold as it results from the determination means 16 can, in contrary to the subsequent description, particularly with regard to Fig. 4, also be a representation by spectral samples of the spectral masking threshold.

The prefilter or preestimation means 18 is coupled to both the masking threshold determination means 16 and the input

12 and filters the output signal for normalizing the same with regard to the masking threshold for obtaining a prefiltered signal  $f(n)$ . The prefilter means 18 is based, for example, on a linear filter and is implemented to  
5 adjust the filter coefficients in dependence on the representation of the masking threshold provided by the masking threshold of the determination means 16, such that the transmission function of the linear filter corresponds substantially to the inverse of the masking threshold.  
10 Adjustment of the filter coefficients can be performed block-wise, half block-wise, such as in the case described below of the blocks overlapping by half in the masking threshold determination, or sample-wise, for example by interpolating the filter coefficients obtained by the  
15 block-wise determined masking threshold representations, or by filter coefficients obtained therefrom across the interblock gaps.

The forward prediction means 20 is coupled to the prefilter means 18, for subjecting the samples  $f(n)$  of the prefiltered signal, which are filtered adaptively in the time domain by using the psychoacoustic masking threshold to a forward-adaptive prediction, for obtaining a predicted signal  $\hat{f}(n)$ , a residual signal  $r(n)$  representing a  
25 prediction error to the prefiltered signal  $f(n)$ , and a representation of prediction filter coefficients, based on which the predicted signal can be reconstructed. Particularly, the forward-adaptive prediction means 20 is implemented to determine the representation of the  
30 prediction filter coefficients immediately from the prefiltered signal  $f$  and not only based on a subsequent quantization of the residual signal  $r$ . Although, as will be discussed in more detail below with reference to Fig. 4, the prediction filter coefficients are represented in the  
35 LFS domain, in particular in the form of a LFS prediction residual, other representations, such as an intermediate representation in the shape of linear filter coefficients, are also possible. Further, means 20 performs the

prediction filter coefficient determination according to the subsequent description exemplarily block-wise, i.e. per block in subsequent block of samples  $f(n)$  of the prefiltered signal, wherein, however, other procedures are also possible. Means 20 is then implemented to determine the predicted signal  $\hat{f}$  via these determined prediction filter coefficients, and to subtract the same from the prefiltered signal  $f$ , wherein the determination of the predicted signal is performed, for example, via a linear filter, whose filter coefficients are adjusted according to the forward-adaptively determined prediction coefficient representations. The residual signal available on the decoder side, i.e. the quantized and clipped residual signal  $i_c(n)$ , added to previously output filter output signal values, can serve as filter input signal, as will be discussed below in more detail.

The quantizing/clip means 22 is coupled to the prediction means 20, for quantizing or clipping, respectively, the residual signal via a quantizing function mapping the values  $r(n)$  of the residual signal to a constant and limited number of quantizing levels, and for transmitting the quantized residual signal obtained in that way in the shape of the quantizing indices  $i_c(n)$ , as has already been mentioned, to the forward-adaptive prediction means 20.

The quantized residual signal  $i_c(n)$ , the representation of the prediction coefficients determined by the means 20, as well as the representation of the masking threshold determined by the means 16 make up information provided to the decoder side via the encoded signal 14, wherein therefore the bit stream generation means 24 is provided exemplarily in Fig. 1, for combining the information according to a serial bit stream or a packet transmission, possibly by using a further lossless encoding.

Before the more detailed structure of the encoder of Fig. 1 will be discussed, the mode of operation of the encoder 1

will be described below based on the above structure of the encoder 10. By filtering the audio signal by the prefilter means 18 with a transmission function corresponding to the inverse of the masking threshold, a prefiltered signal  $f(n)$  results, which obtains a spectral power density of the error by uniform quantizing, which mainly corresponds to a white noise, and would result in a noise spectrum similar to the masking threshold by filtering in the postfilter on the decoder side. However, first, the residual signal  $f$  is reduced to a prediction error  $r$  by the forward-adaptive prediction means 20 by a forward adapted predicted signal  $\hat{f}$  by subtraction. The subsequent coarse quantization of this prediction error  $r$  by the quantizing/clipping means 22 has no effect on the prediction coefficients of the prediction means 20, neither on the encoder nor the decoder side, since the calculation of the prediction coefficients is performed in a forward-adaptive manner and thus based on the unquantized values  $f(n)$ . Quantization is not only performed in a coarse way, in the sense that a coarse quantizing step size is used, but is also performed in a coarse manner in the sense that even quantization is performed only to a constant and limited number of quantizing levels, so that for representing every quantized residual signal  $i_c(n)$  or every quantizing index in the encoded audio signal 14 only a fixed number of bits is required, which allows inherently a constant bit rate with regard to the residual values  $i_c(n)$ . As will be described below, quantization is performed mainly by quantizing to uniformly spaced quantizing levels of fixed number, and below exemplarily to a number of a merely three quantizing levels, wherein quantization is performed, for example, such that an unquantized residual signal value  $r(n)$  is quantized to the next quantizing level, for obtaining the quantizing index  $i_c(n)$  of the corresponding quantizing level for the same. Extremely high and extremely low values of the unquantized residual signal  $r(n)$  are thus mapped to the respective highest or lowest, respectively, quantizing level or the respective quantizing level index,

respectively, even when they would be mapped to a higher quantizing level at uniform quantizing with the same step size. In so far, the residual signal  $r$  is also "clipped" or limited, respectively, by the means 22. However, the latter  
5 has the effect, as will be discussed below, that the error PSD (PSD = power spectral density) of the prefiltered signal is no longer a white noise, but is approximated to the signal PSD of the prefiltered signal depending on the degree of clipping. On the decoder side, this has the  
10 effect that the noise PSD remains below the signal PSD even at bit rates that are lower than predetermined by the masking threshold.

In the following, the structure of the encoder in Fig. 1  
15 will be described in more detail. Particularly, the masking threshold determination means 16 comprises a masking threshold determiner or a perceptual model 26, respectively, operating according to the perceptual model, a prefilter coefficient calculation module 28 and a  
20 coefficient encoder 30, which are connected in the named order between the input 12 and the prefilter means 18 as well as the bit stream generator 24. The prefilter means 18 comprises a coefficient decoder 32 whose input is connected to the output of the coefficient encoder 30, as well as the  
25 prefilter 34, which is, for example, an adaptive linear filter, and which is connected with its data input to the input 12 and with its data output to the means 20, while its adaption input for adapting the filter coefficients is connected to an output of the coefficient decoder 32. The  
30 prediction means 20 comprises a prediction coefficient calculation module 36, a coefficient encoder 38, a coefficient decoder 40, a subtractor 42, a prediction filter 44, a delay element 46, a further adder 48 and a dequantizer 50. The prediction coefficient calculation  
35 module 46 and the coefficient encoder 38 are connected in series in this order between the output of the prefilter 34 and the input of the coefficient decoder 40 or a further input of the bit stream generator 24, respectively, and

cooperate for determining a representation of the prediction coefficients block-wise in a forward-adaptive manner. The coefficient decoder 40 is connected between the coefficient encoder 38 and the prediction filter 44, which is, for example, a linear prediction filter. Apart from the prediction coefficient input connected to the coefficient decoder 40, the filter 44 comprises a data input and a data output, to which the same is connected in a closed loop, which comprises, apart from the filter 44, the adder 48 and the delay element 46. Particularly, the delay element 46 is connected between the adder 48 and the filter 44, while the data output of the filter 44 is connected to a first input of the adder 48. Above that, the data output of the filter 44 is also connected to an inverting input of the subtractor 42. A non-inverting input of the subtractor 42 is connected to the output of the prefilter 34, while the second input of the adder 48 is connected to an output of the dequantizer 50. A data input of the dequantizer 50 is coupled to the quantizing/clipping means 22 as well as to a step size control input of the dequantizer 50. The quantizing/clipping means 22 comprises a quantizer module 52 as well as a step size adaption block 54, wherein again the quantizing module 52 consists of a uniform quantizer 56 with uniform and controllable step size and a limiter 58, which are connected in series in the named order between an output of the subtractor 42 and the further input of the bit stream generator 24, and wherein the step size adaption block 54 again comprises a step size adaption module 60 and a delay member 62, which are connected in series in the named order between the output of the limiter 58 and a step size control input of the quantizer 56. Additionally, the output of the limiter 58 is connected to the data input of the dequantizer 50, wherein the step size control input of the dequantizer 50 is also connected to the step size adaption block 60. An output of the bit stream generator 24 again forms the output 14 of the encoder 10.

After the detailed structure of the encoder of Fig. 1 has been described in detail above, its mode of operation will be described below. The perceptual model module 26 determines or estimates, respectively, the masking threshold in a block-wise manner from the audio signal. Therefore, the perceptual model module 26 uses, for example, a DFT of the length 256, i.e. a block length of 256 samples  $x(n)$ , with 50% overlapping between the blocks, which results in a delay of the encoder 10 of 128 samples of the audio signal. The estimation of the masking threshold output by the perceptual model module 26 is, for example, represented in a spectrally sampled form in a Bark band or linear frequency scale. The masking threshold output per block by the perceptual model module 26 is used in the coefficient calculation module 24 for calculating filter coefficients of a predetermined filter, namely the filter 34. The coefficients calculated by the module 28 can, for example, be LPC coefficients, which model the masking threshold. The prefilter coefficients for every block are again encoded by the coefficient encoder 30, which will be discussed in more detail with reference to Fig. 4. The coefficient decoder 34 decodes the encoded prefilter coefficients for retrieving the prefilter coefficients of the module 28, wherein the prefilter 34 again obtains these parameters or prefilter coefficients, respectively, and uses the same, so that it normalizes the input signal  $x(n)$  with regard to its masking threshold or filters the same with a transmission function, respectively, which essentially corresponds to the inverse of the masking threshold. Compared to the input signal, the resulting prefiltered signal  $f(n)$  is significantly smaller in amount.

In the prediction coefficient calculation module 36, the samples  $f(n)$  of the prefiltered signal are processed in a block-wise manner, wherein the block-wise division can correspond exemplarily to the one of the audio signal 12 by the perceptual model module 26, but does not have to do

this. For every block of prefiltered samples, the coefficient calculation module 36 calculates prediction coefficients for usage by the prediction filter 44. Therefore, the coefficient calculation module 36 performs, for example, LPC (LPC = linear predictive coding) analysis per block of the prefiltered signal for obtaining the prediction coefficients. The coefficient encoder 38 encodes then the prediction coefficients similar to the coefficient encoder 30, as will be discussed in more detail below, and outputs this representation of the prediction coefficients to the bit stream generator 24 and particularly the coefficient decoder 40, wherein the latter uses the obtained prediction coefficient representation for applying the prediction coefficients obtained in the LPC analysis by the coefficient calculation module 36 to the linear filter 44, so that the closed loop predictor consisting of the closed loop of filter 44, delay member 46 and adder 48 generates the predicted signal  $\hat{f}(n)$ , which is again subtracted from the prefiltered signal  $f(n)$  by the subtractor 42. The linear filter 44 is, for example, a linear prediction filter of the type  $A(z) = \sum_{i=1}^N a_i z^{-i}$  of the length  $N$ , wherein the coefficient decoder 40 adjusts the values  $a_i$  in dependence on the prediction coefficients calculated by the coefficient calculation module 36, i.e. the weightings with which the previous predicted values  $\hat{f}(n)$  plus the dequantized residual signal values are weighted and then summed for obtaining the new or current, respectively, predicted value  $\hat{f}$ .

The prediction remainder  $r(n)$  obtained by the subtractor 22 is subject to uniform quantization, i.e. quantization with uniform quantizing step size, in the quantizer 56, wherein the step size  $\Delta(n)$  is time-variable, and is calculated or determined, respectively, by the step size adaption module in a backward-adaptive manner, i.e. from the quantized residual values to the previous residual values  $r(m < n)$ . More precisely, the uniform quantizer 56 outputs a quantized residual value  $q(n)$  per residual value  $r(n)$ ,

which can be expressed as  $q(n) = i(n) \cdot \Delta(n)$  and can be referred to as provisional quantizing step with index. The provisional quantizing index  $i(n)$  is again clipped by the limiter 58, to the amount  $C = [-c;c]$ , wherein  $c$  is a  
5 constant  $c \in \{1,2,\dots\}$ . Particularly, the limiter 58 is implemented such that all provisional index values  $i(n)$  with  $|i(n)| > c$  are either set to  $-c$  or  $c$ , depending on which is closer. Merely the clipped or limited, respectively, index sequence or series  $i_c(n)$  is output by the limiter 58  
10 to the bit stream generator 24, the dequantizer 50 and the step size adaption block 54 or the delay element 62, respectively, because the delay member 62, as well as all other delay members in the present embodiments, delays the incoming values by one sample.

15 Now, backward-adaptive step size control is realized via the step size adaption block 54, in that the same uses past index sequence values  $i_c(n)$  delayed by the delay member 62 for constantly adapting the step size  $\Delta(n)$ , such that the area limited by the limiter 58, i.e. the area set by the  
20 "allowed" quantizing indices or the corresponding quantizing levels, respectively, is placed such to the statistic probability of occurrence of unquantized residual values  $r(n)$ , that the allowed quantizing levels occur as  
25 uniformly as possible in the generated clipped quantizing index sequence stream  $i_c(n)$ . Particularly, the step size adaption module 60 calculates, for example, the current step size  $\Delta(n)$  for example by using the two immediately preceding clipped quantizing indices  $i_c(n-1)$  and  $i_c(n-2)$  as  
30 well as the immediately previously determined step size value  $\Delta(n-1)$  to  $\Delta(n) = \beta \Delta(n-1) + \delta(n)$ , with  $\beta \in [0.0;1.0[$ ,  $\delta(n) = \delta_0$  for  $|i_c(n-1) + i_c(n-2)| \leq I$  and  $\delta(n) = \delta_1$  for  $|i_c(n-1) + i_c(n-2)| > I$ , wherein  $\delta_0$ ,  $\delta_1$  and  $I$  are appropriately adjusted constants, as well as  $\beta$ .

35 As will be discussed in more detail below with reference to Fig. 5, the decoder uses the obtained quantizing index sequence  $i_c(n)$  and the step size sequence  $\Delta(n)$ , which is

also calculated in a backward-adaptive manner for reconstructing the dequantized residual value sequence  $q_c(n)$  by calculating  $i_c(n) \cdot \Delta(n)$ , which is also performed in the encoder 10 of Fig. 1, namely by the dequantizer 50 in the prediction means 20. Like on the decoder side, the residual value sequence  $q_c(n)$  constructed in that way is subject to an addition with the predicted values  $\hat{f}(n)$  in a sample-wise manner, wherein the addition is performed in the encoder 10 via the adder 48. While the reconstructed or dequantized, respectively, prefiltered signal obtained in that way is no longer used in the encoder 10, except for calculating the subsequent predicted values  $\hat{f}(n)$ , the postfilter generates the decoded audio sample sequence  $y(n)$  therefrom on the decoder side, which cancels the normalization by the prefilter 34.

The quantizing noise introduced in the quantizing index sequence  $q_c(n)$  is no longer white due to the clipping. Rather, its spectral form copies the one of the prefiltered signal. For illustrating this, reference is briefly made to Fig. 3, which shows, in graphs a, b and c, the PSD of the prefiltered signal (upper graph) and the PSD of the quantizing error (respective lower graph) for different numbers of quantizing levels or stages, respectively, namely for  $C = [-15;15]$  in graph a, for a limiter range of  $[-7;7]$  in graph b, and a clipping range of  $[-1;1]$  in graph c. For clarity reasons, it should further be noted that the PSD courses of the error PSDs in graphs A-C have each been plotted with an offset of -10dB. As can be seen, the prefiltered signal corresponds to a colored noise with a power of  $\sigma^2 = 34$ . At a quantization with a step size  $\Delta = 1$ , the signal lies within  $[-21;21]$ , i.e. the samples of the prefiltered signal have an occurrence distribution or form a histogram, respectively, which lies within this domain. For graphs a to c in Fig. 3, the quantizing range has been limited, as mentioned, to  $[-15;15]$  in a),  $[-7;7]$  in b) and  $[-1;1]$  in c). The quantizing error has been measured as the difference between the unquantized prefiltered signal and

the decoded prefiltered signal. As can be seen, a quantizing noise is added to the prefiltered signal by increasing clipping or with increasing limitation of the number of quantizing levels, which copies the PSD of the prefiltered signal, wherein the degree of copying depends on the hardness or the extension, respectively, of the applied clipping. Consequently, after postfiltering, the quantizing noise spectrum on the decoder side copies more the PSD of the audio input signal. This means that the quantizing noise remains below the signal spectrum after decoding. This effect is illustrated in Fig. 2, which shows in graph a, for the case of backward-adaptive prediction, i.e. prediction according to the above described comparison ULD scheme, and in graph b, for the case of forward-adaptive prediction with applied clipping according to Fig. 1, respectively three courses in a normalized frequency domain, namely, from top to bottom, the signal PSD, i.e. the PSD of the audio signal, the quantizing error PSD or the quantizing noise after decoding (straight line) and the masking threshold (dotted line). As can be seen, the quantizing noise for the comparison ULD encoder (Fig. 2a) is formed like the masking threshold and exceeds the signal spectrum for portions of the signal. The effect of the forward-adaptive prediction of the prefiltered signal combined with subsequent clipping or limiting, respectively, of the quantizing level number is now clearly illustrated in Fig. 2b, where it can be seen that the quantizing noise is always lower than the signal spectrum and its shape represents a mixture of the signal spectrum and the masking threshold. In listening tests, it has been found out that the encoding artifacts according to Fig. 2b are less spurious, i.e. the perceived listening quality is better.

The above description of the mode of operation of the encoder of Fig. 1 concentrated on the postprocessing of the prefiltered signal  $f(n)$ , for obtaining the clipped quantizing indices  $i_c(n)$  to be transmitted to the decoder

side. Since they originate from an amount with a constant and limited number of indices, they can each be represented with the same number of bits within the encoded data stream at the output 14. Therefore, the bit stream generator 24  
5 uses, for example, an injective mapping of the quantizing indices to  $m$  bit words that can be represented by a predetermined number of bits  $m$ .

The following description deals with the transmission of  
10 the prefilter or prediction coefficients, respectively, calculated by the coefficient calculation modules 28 and 36 to the decoder side, i.e. particularly with an embodiment for the structure of the coefficient encoders 30 and 38.

15 As is shown, the coefficient encoders according to the embodiment of Fig. 4 comprise an LSF conversion module 102, a first subtractor 104, a second subtractor 106, a uniform quantizer 108 with uniform and adjustable quantizing step size, a limiter 110, a dequantizer 112, a third adder 114,  
20 two delay members 116 and 118, a prediction filter 120 with fixed filter coefficients or constant filter coefficients, respectively, as well as a step size adaption module 122. The filter coefficients to be encoded come in at an input 124, wherein an output 126 is provided for outputting the  
25 encoded representation.

An input of the LSF conversion module 102 directly follows the input 124. The subtractor 104 with its non-inverting input and its output is connected between the output of the  
30 LSF conversion module 102 and a first input of the subtractor 106, wherein a constant  $l_c$  is applied to the input of the subtractor 104. The subtractor 106 is connected with its non-inverting input and its output between the first subtractor 104 and the quantizer 108,  
35 wherein its inverting input is coupled to an output of the prediction filter 120. Together with the delay member 118 and the adder 114, the prediction filter 120 forms a closed-loop predictor, in which the same are connected in

series in a loop with feedback, such that the delay member 118 is connected between the output of the adder 114 and the input of the prediction filter 120, and the output of the prediction filter 120 is connected to a first input of the adder 114. The remaining structure corresponds again mainly to the one of the means 22 of the encoder 10, i.e. the quantizer 108 is connected between the output of the subtractor 106 and the input of the limiter 110, whose output is again connected to the output 126, an input of the delay member 116 and an input of the dequantizer 112. The output of the delay member 116 is connected to an input of the step size adaption module 122, which thus form together a step size adaption block. An output of the step size adaption module 122 is connected to step size control inputs of the quantizer 108 and the dequantizer 112. The output of the dequantizer 112 is connected to the second input of the adder 114.

After the structure of the coefficient encoder has been described above, its mode of operation will be described below, wherein reference is made again to Fig. 1. The transmission of both the prefilters and the prediction or predictor coefficients, respectively, or their encoding, respectively, is performed by using a constant bit rate encoding scheme, which is realized by the structure according to Fig. 4. Then, in the LSF conversion module 102, the filter coefficients, i.e. the prefilter or prediction coefficients, respectively, are first converted to LSF values  $l(n)$  or transferred to the LSF domain, respectively. Every spectral line frequency  $l(n)$  is then processed by the residual elements in Fig. 4 as follows. This means the following description relates to merely one spectral line frequency, wherein the processing of course, is performed for all spectral line frequencies. For example, the module 102 generates LSF values for every set of prefilter coefficients representing a masking threshold, or a block of prediction coefficients predicting the prefiltered signal. The subtractor 104 subtracts a constant

reference value  $l_c$  from the calculated value  $l(n)$ , wherein a sufficient range for  $l_c$  ranges, for example, from 0 to  $\pi$ . From the resulting difference  $l_d(n)$ , the subtractor 106 subtracts a predicted value  $\hat{l}_d(n)$ , which is calculated by  
5 the closed-loop predictor 120, 118 and 114 including the prediction filter 120, such as a linear filter, with fixed coefficients  $A(z)$ . What remains, i.e. the residual value, is quantized by the adaptive step size quantizer 108, wherein the quantizing indices output by the quantizer 108  
10 are clipped by the limiter 110 to a subset of the quantizing indices received by the same, such as, for example, that for all clipped quantizing indices  $l_e(n)$ , as they are output by the limiter 110, the following applies:  $\forall : l_e(n) \in \{-1, 0, 1\}$ . For quantizing step size adaption of  
15  $\Delta(n)$  of the LSF residual quantizer 108, the step size adaption module 122 and the delay member 116 cooperate for example in the way described with regard to the step size adaption block 54 with reference to Fig. 1, however, possibly with a different adaption function or with  
20 different constants  $\beta$ ,  $I$ ,  $\delta_0$ ,  $\delta_1$  and  $I$ . While the quantizer 108 uses the current step size for quantizing the current residual value to  $l_e(n)$ , the dequantizer 112 uses the step size  $\Delta_1(n)$  for dequantizing this index value  $l_e(n)$  again and for supplying the resulting reconstructed value  
25 for the LSF residual value, as it has been output by the subtractor 106, to the adder 114, which adds this value to the corresponding predicted value  $\hat{l}_d(n)$ , and supplies the same via the delay member 118 delayed by a sample to the filter 120 for calculating the predicted LSF value  $\hat{l}_d(n)$   
30 for the next LSF value  $l_d(n)$ .

If the two coefficient encoders 30 and 38 are implemented in the way described in Fig. 4, the coder 10 of Fig. 1 fulfills a constant bit rate condition without using any  
35 loop. Due to the block-wise forward adaption of the LPC coefficients and the applied encoding scheme, no explicit reset of the predictor is required.

Before results of listening tests, which have been obtained by an encoder according to Figs. 1 and 4, will be discussed below, the structure of a decoder according to an embodiment of the present invention will be described below, which is suitable for decoding an encoded data stream from this encoder, wherein reference is made to Figs. 5 and 6. Fig. 6 also shows the structure of the coefficient decoder in Fig. 1.

10 The decoder generally indicated by 200 in Fig. 5 comprises an input 202 for receiving the encoded data stream, an output 204 for outputting the decoded audio stream  $y(n)$  as well as a dequantizing means 206 having a limited and constant number of quantizing levels, a prediction means 15 208, a reconstruction means 210 as well as a postfilter means 212. Additionally, an extractor 214 is provided, which is coupled to the input 202 and implemented to extract, from the incoming encoded bit stream, the quantized and clipped prefilter residual signal  $i_c(n)$ , the 20 encoded information about the prefilter coefficients and the encoded information about the prediction coefficients, as they have been generated from the coefficient encoders 30 and 38 (Fig. 1) and to output the same at the respective outputs. The dequantizing means 206 is coupled to the 25 extractor 214 for obtaining the quantizing indices  $i_c(n)$  from the same and for performing dequantization of these indices to a limited and constant number of quantizing levels, namely - sticking to the same notation as above -  $\{-c \cdot \Delta(n); c \cdot \Delta(n)\}$ , for obtaining a dequantized or 30 reconstructed prefilter signal  $q_c(n)$ , respectively. The prediction means 208 is coupled to the extractor 214 for obtaining a predicted signal for the prefiltered signal, namely  $\hat{f}_{c(n)}$  from the information about the prediction coefficients. The prediction means 208 is coupled to the 35 extractor 214 for determining a predicted signal for the prefiltered signal, namely  $\hat{f}(n)$ , from the information about the prediction coefficients, wherein the prediction means 208 according to the embodiment of Fig. 5 is also

connected to an output of the reconstruction means 210. The reconstruction means 210 is provided for reconstructing the prefiltered signal, based on the predicted signal  $\hat{f}(n)$  and the dequantized residual signals  $q_c(n)$ . This reconstruction is then used by the subsequent postfilter means 212 for filtering the prefiltered signal based on the prefilter coefficient information received from the extractor 214, such that the normalization with regard to the masking threshold is canceled for obtaining the decoded audio signal  $y(n)$ .

After the basic structure of the decoder of Fig. 5 has been described above, the structure of the decoder 200 will be discussed in more detail. Particularly, the dequantizer 206 comprises a step size adaption block of a delay member 216 and a step size adaption module 218 as well as a uniform dequantizer 220. The dequantizer 220 is connected to an output of the extractor 214 with its data input, for obtaining the quantizing indices  $i_c(n)$ . Further, the step size adaption module 218 is connected to this output of the extractor 214 via the delay member 216, whose output is again connected to a step size control input of the dequantizer 220. The output of the dequantizer 220 is connected to a first input of the adder 222, which forms the reconstruction means 210. The prediction means 208 comprises a coefficient decoder 224, a prediction filter 226 as well as delay member 228. Coefficient decoder 224, adder 222, prediction filter 226 and delay member 228 correspond to elements 40, 44, 46 and 48 of the encoder 10 with regard to their mode of operation and their connectivity. In particular, the output of the prediction filter 226 is connected to the further input of the adder 222, whose output is again fed back to the data input of the prediction filter 226 via the delay member 228, as well as coupled to the postfilter means 212. The coefficient decoder 224 is connected between a further output of the extractor 214 and the adaption input of the prediction filter 226. The postfilter means comprises a coefficient

decoder 230 and a postfilter 232, wherein a data input of the postfilter 232 is connected to an output of the adder 222 and a data output of the postfilter 232 is connected to the output 204, while an adaption input of the postfilter 232 is connected to an output of the coefficient decoder 230 for adapting the postfilter 232, whose input again is connected to a further output of the extractor 214.

10 As has already been mentioned, the extractor 214 extracts the quantizing indices  $i_c(n)$  representing the quantized prefilter residual signal from the encoded data stream at the input 202. In the uniform dequantizer 220, these quantizing indices are dequantized to the quantized residual values  $q_c(n)$ . Inherently, this dequantizing remains within the allowed quantizing levels, since the quantizing indices  $i_c(n)$  have already been clipped on the encoder side. The step size adaption is performed in a backward-adaptive manner, in the same way as in the step size adaption block 54 of the encoder of Fig. 1. Without transmission errors, the dequantizer 220 generates the same values as the dequantizer 50 of the encoder of Fig. 1. Therefore, the elements 222, 226, 228 and 224 based on the encoded prediction coefficients obtain the same result as it is obtained in the encoder 10 of Fig. 1 at the output of the adder 48, i.e. a dequantized or reconstructed prefilter signal, respectively. The latter is filtered in the postfilter 232, with a transmission function corresponding to the masking threshold, wherein the postfilter 232 is adjusted adaptively by the coefficient decoder 230, which appropriately adjust the postfilter 230 or its filter coefficients, respectively, based on the prefilter coefficient information.

35 Assuming that the encoder 10 is provided with coefficient encoders 30 and 38, which are implemented as described in Fig. 4, the coefficient decoders 224 and 230 of the encoder 200 but also the coefficient decoder 40 of the encoder 10

are structured as shown in Fig. 6. As can be seen, a coefficient decoder comprises two delay members 302, 304, a step size adaption module 306 forming a step size adaption block together with the delay member 302, a uniform  
5 dequantizer 308 with uniform step size, a prediction filter 310, two adders 312 and 314, an LSF reconversion module 316 as well as an input 318 for receiving the quantized LSF residual values  $l_e(n)$  with constant offset  $-l_c$  and an output 320 for outputting the reconstructed prediction or  
10 prefilter coefficients, respectively. Thereby, the delay member 302 is connected between an input of the step size adaption module 306 and the input 318, an input of the dequantizer 308 is also connected to the input 318, and a step size adaption input of the dequantizer 308 is  
15 connected to an output of the step size adaption module 306. The mode of operation and connectivity of the elements 302, 306 and 308 corresponds to the one of 112, 116 and 122 in Fig. 4. A closed-loop predictor of delay member 304, prediction filter 310 and adder 312, which are  
20 connected in a common loop by connecting the delay member 304 between an output of the adder 312 and an input of the prediction filter 310, and by connecting a first input of the adder 312 to the output of the dequantizer 308, and by connecting a second input of the adder 312 to an output of  
25 the prediction filter 310, is connected to an output of the dequantizer 308. Elements 304, 310 and 312 correspond to the elements 120, 118 and 114 of Fig. 4 in their mode of operation and connectivity. Additionally, the output of the adder 312 is connected to a first input of the adder 314,  
30 at the second input of which the constant value  $l_c$  is applied, wherein, according to the present embodiment, the constant  $l_c$  is an agreed amount, which is present to both encoder and the decoder and thus does not have to be transmitted as part of the side information, although the  
35 latter would also be possible. The LSF reconversion module 316 is connected between an output of the adder 314 and the output 320.

The LSF residual signal indices  $l_e(n)$  incoming at the input 318 are dequantized by the dequantizer 308, wherein the dequantizer 308 uses the backward-adaptive step size values  $\Delta(n)$ , which had been determined in a backward-adaptive manner by the step size adaption module 306 from already dequantized quantizing indices, namely those that had been delayed by a sample by the delay member 302. The adder 312 adds the predicted signal to the dequantized LSF residual values, which calculates the combination of delay member 304 and prediction filter 210 from sums that the adder 312 has already calculated previously and thus represent the reconstructed LSF values, which are merely provided with a constant offset by the constant offset  $l_c$ . The latter is corrected by the adder 314 by adding the value  $l_c$  to the LSF values, which the adder 312 outputs. Thus, at the output of the adder 314, the reconstructed LSF values result, which are converted by the module 316 from the LSF domain back to reconstructed prediction or prefilter coefficients, respectively. Therefore, the LSF reconversion module 316 considers all spectral line frequencies, whereas the discussion of the other elements of Fig. 6 was limited to the description of one spectral line frequency. However, the elements 302-314 perform the above-described measures also at the other spectral line frequencies.

After providing both encoder and decoder embodiments above, listening test results will be presented below based on Fig. 7, as they have been obtained via an encoding scheme according to Figs. 1, 4, 5 and 6. In the performed tests, both an encoder according to Figs. 1, 4 and 6 and an encoder according to the comparison ULD encoding scheme discussed at the beginning of the description of the Figs. have been tested, in a listening test according to the MUSHRA standard, where the moderators have been omitted. The MUSHRA test has been performed on a laptop computer with external digital-to-analog converter and STAX amplifier/headphones in a quiet office environment. The

group of eight test listeners was made up of expert and non-expert listeners. Before the participants began the listening test, they had the opportunity to listen to a test set. The tests have been performed with twelve mono  
5 audio files of the MPEG test set, wherein all had a sample frequency of 32 kHz, namely es01 (Suzanne Vega), es02 (male speech, German), es03 (female speech, English), sc01 (trumpet), sc02 (orchestra), sc03 (pop music), si01 (cembalo), si02 (castanets), si03 (pitch pipe), sm01  
10 (bagpipe), sm02 (glockenspiel), sm03 (puckled strings).

For the comparison ULD encoding scheme, a backward-adaptive prediction with a length of 64 has been used in the implementation, together with a backward-adaptive Golomb  
15 encoder for entropy encoding, with a constant bit rate of 64 kBit/s. In contrast, for implementing the encoder according to Figs. 1, 4 and 6, a forward-adaptive predictor with a length of 12 has been used, wherein the number of different quantizing levels has been limited to 3, namely  
20 such that  $\forall n : i_c(n) \in \{-1, 0, 1\}$ . This resulted, together with the encoded side information, in a constant bit rate of 64 kBit/s, which means the same bit rate.

The results of the MUSHRA listening tests are shown in  
25 Fig. 7, wherein both the average values and 95 % confidence intervals are shown, for the twelve test pieces individually and for the overall result across all pieces. As long as the confidence intervals overlap, there is no statistically significant difference between the encoding  
30 methods.

The piece es01 (Suzanne Vega) is a good example for the superiority of the encoding scheme according to Figs. 1, 4, 5 and 6 at lower bit rates. The higher portions of the  
35 decoded signal spectrum show less audible artifacts compared to the comparison ULD encoding scheme. This results in a significantly higher rating of the scheme according to Figs. 1, 4, 5 and 6.

The signal transients of the piece sm02 (Glockenspiel) have a high bit rate requirement for the comparison ULD encoding scheme. In the used 64kBit/s, the comparison ULD encoding scheme generates spurious encoding artifacts across full blocks of samples. In contrast, the encoder operating according to Figs. 1, 4 and 6 provides a significantly improved listening quality or perceptual quality, respectively. The overall rating, seen in the graph of Fig. 7 on the right, of the encoding scheme formed according to Figs. 1, 4 and 6 obtained a significantly better rating than the comparison ULD encoding scheme. Overall, this encoding scheme got an overall rating of "good audio quality" under the given test conditions.

In summary, from the above-described embodiments, an audio encoding scheme with low delay results, which uses a block-wise forward-adaptive prediction together with clipping/limiting instead of a backward-adaptive sample-wise prediction. The noise shaping differs from the comparison ULD encoding scheme. The listening test has shown that the above-described embodiments are superior to the backward-adaptive method according to the comparison ULD encoding scheme in the case of lower bit rates. Subsequently, the same are a candidate for closing the bit rate gap between high quality voice encoders and audio encoders with low delay. Overall, the above-described embodiments provided a possibility for audio encoding schemes having a very low delay of 6 - 8 ms for reduced bit rates, which has the following advantages compared to the comparison ULD encoder. The same is more robust against high quantizing errors, has additional noise shaping abilities, has a better ability for obtaining a constant bit rate, and shows a better error recovery behavior. The problem of audible quantizing noise at positions without signal, as is the case in the comparison ULD encoding scheme, is addressed by the embodiment by a modified way of increasing the quantizing noise above the masking

threshold, namely by adding the signal spectrum to the masking threshold instead of uniformly increasing the masking threshold to a certain degree. In that way, there is no audible quantizing noise at positions without signal.

5

In other words, the above embodiments differ from the comparison ULD encoding scheme in the following way. In the comparison ULD encoding scheme, backward-adaptive prediction is used, which means that the coefficients for the prediction filter  $A(z)$  are updated on a sample-by-sample basis from previously decoded signal values. A quantizer having a variable step size is used, wherein the step size adapts all 128 samples by using information from the entropy encoders and the same is transmitted as side information to the decoder side. By this procedure, the quantizing step size is increased, which adds more white noise to the prefiltered signal and thus uniformly increases the masking threshold. If the backward-adaptive prediction is replaced with a forward-adaptive block-wise prediction in the comparison ULD encoding scheme, which means that the coefficients for the prediction filter  $A(z)$  are calculated once for 128 samples from the unquantized prefiltered samples, and transmitted as side information, and if the quantizing step size is adapted for the 128 samples by using information from the entropy encoder and transmitted as side information to the decoder side, the quantizing step size is still increased, as it is the case in the comparison ULD encoding scheme, but the predictor update is unaffected by any quantization. The above embodiments used only a forward adapted block-wise prediction, wherein additionally the quantizer had merely a given number  $2N+1$  of quantizing stages having a fixed step size. For the prefiltered signals  $x(n)$  with amplitudes outside the quantizer range  $[-N\Delta; N\Delta]$  the quantized signal was limited to  $[-N\Delta; N\Delta]$ . This results in a quantizing noise having a PSD, which is no longer white, but copies the PSD of the input signal, i.e. the prefiltered audio signal.

10

15

20

25

30

35

As a conclusion, the following is to be noted on the above embodiments. First, it should be noted that different possibilities exist for transmitting information about the representation of the masking threshold, as they are obtained by the perceptual model module 26 within the encoder to the prefilter 34 or prediction filter 44, respectively, and to the decoder, and there particularly to the postfilter 232 and the prediction filter 226. Particularly, it should be noted that it is not required that the coefficient decoders 32 and 40 within the encoder receive exactly the same information with regard to the masking threshold, as it is output at the output 14 of the encoder and as it is received at the output 202 of the decoder. Rather, it is possible, that, for example in a structure of the coefficient encoder 30 according to Fig. 4, the obtained indices  $l_e(n)$  as well as the prefilter residual signal quantizing indices  $i_c(n)$  originate also only from an amount of three values, namely -1, 0, 1, and that the bit stream generator 24 maps these indices just as clearly to corresponding  $n$  bit words. According to an embodiment according to Figs. 1, 4 or 5, 6, respectively, the prefilter quantizing indices, the prediction coefficient quantizing indices and/or the prefilter quantizing indices each originating from the amount -1, 0, 1, are mapped in groups of fives to a 8-bit word, which corresponds to a mapping of  $3^5$  possibilities to  $2^8$  bit words. Since the mapping is not surjective, several 8-bit words remain unused and can be used in other ways, such as for synchronization or the same.

On this occasion, the following should be noted. Above, it has been described with reference to Fig. 6 that the structure of the coefficient decoders 32 and 230 is identical. In this case, the prefilter 34 and the postfilter 232 are implemented such that when applying the same filter coefficients they have a transmission function inverse to each other. However, it is of course also

possible that, for example, the coefficient encoder 32 performs an additional conversion of the filter coefficients, so that the prefilter has a transmission function mainly corresponding to the inverse of the masking  
5 threshold, whereas the postfilter has a transmission function mainly corresponding to the masking threshold.

In the above embodiments, it has been assumed that the masking threshold is calculated in the module 26. However,  
10 it should be noted that the calculated threshold does not have to exactly correspond to the psychoacoustic threshold, but can represent a more or less exact estimation of the same, which might not consider all psychoacoustic effects but merely some of them. Particularly, the threshold can  
15 represent a psychoacoustically motivated threshold, which has been deliberately subject to a modification in contrast to an estimation of the psychoacoustic masking threshold.

Further, it should be noted that the backward-adaptive  
20 adaption of the step size in quantizing the prefilter residual signal values does not necessarily have to be present. Rather, in certain application cases, a fixed step size can be sufficient.

25 Further, it should be noted that the present invention is not limited to the field of audio encoding. Rather, the signal to be encoded can also be a signal used for stimulating a fingertip in a cyber-space glove, wherein the perceptual model 26 in this case considers certain tactile  
30 characteristics, which the human sense of touch can no longer perceive. Another example for an information signal to be encoded would be, for example, a video signal. Particularly the information signal to be encoded could be a brightness information of a pixel or image point,  
35 respectively, wherein the perceptual model 26 could also consider different temporal, local and frequency psychovisual covering effects, i.e. a visual masking threshold.

Additionally, it should be noted that quantizer 56 and limiter 58 or quantizer 108 and limiter 110, respectively, do not have to be separate components. Rather, the mapping of the unquantized values to the quantized/clipped values could also be performed by a single mapping. On the other hand, the quantizer 56 or the quantizer 108, respectively, could also be realized by a series connection of a divider followed by a quantizer with uniform and constant step size, where the divider would use the step size value  $\Delta(n)$  obtained from the respective step size adaption module as divisor, while the residual signal to be encoded formed the dividend. The quantizer having a constant and uniform step size could be provided as simple rounding module, which rounds the division result to the next integer, whereupon the subsequent limiter would then limit the integer as described above to an integer of the allowed amount C. In the respective dequantizer, a uniform dequantization would simply be performed with  $\Delta(n)$  as multiplicator.

Further, it should be noted that the above embodiments were restricted to applications having a constant bit rate. However, the present invention is not limited thereto and thus quantization by clipping of, for example, the prefiltered signal used in these embodiments is only one possible alternative. Instead of clipping, a quantizing function with nonlinear characteristic curve could be used. For illustrating this, reference is made to Figs. 8a to 8c. Fig. 8a shows the above-used quantizing function resulting in clipping on three quantizing stages, i.e. a step function with three stages 402a, b, c, which maps unquantized values (x axis) to quantizing indices (y axis), wherein the quantizing stage height or quantizing step size  $\Delta(n)$  is also marked. As can be seen, unquantized values higher than  $\Delta(n)/2$  are clipped to the respective next stage 402a or c, respectively. Fig. 8b shows generally a quantizing function resulting in clipping to  $2n+1$  quantizing stages. The quantizing step size  $\Delta(n)$  is again

shown. The quantizing functions of Figs. 8a and 8b represent quantizing functions, where the quantization between thresholds  $-\Delta(n)$  and  $\Delta(n)$  or  $-\Delta(n)$  and  $\Delta(n)$  takes place in uniform manner, i.e. with the same stage height, whereupon the quantizing stage function proceeds in a flat way, which corresponds to clipping. Fig. 8c shows a nonlinear quantizing function, where the quantizing function proceeds across the area between  $-\Delta(n)$  and  $\Delta(n)$  not completely flat but with a lower slope, i.e. with a larger step size or stage height, respectively, compared to the first area. This nonlinear quantization does not inherently result in a constant bit rate, as it was the case in the above embodiments, but also generates the above-described deformation of the quantizing noise, so that the same adjusts to the signal PSD. Merely as a precautionary measure, it should be noted with reference to Figs. 8a-c, that instead of the uniform quantizing areas non-uniform quantization could be used, where, for example, the stage height increases continuously, wherein the stage heights could be scalable via a stage height adjustment value  $\Delta(n)$  while maintaining their mutual relations. Therefore, for example, the unquantized value could be mapped via a nonlinear function to an intermediate value in the respective quantizer, wherein either before or afterwards multiplication with  $\Delta(n)$  is performed, and finally the resulting value is uniformly quantized. In the respective dequantizer, the inverse would be performed, which means uniform dequantization via  $\Delta(n)$  followed by inverse nonlinear mapping or, conversely, nonlinear conversion mapping at first followed by dequantization with  $\Delta(n)$ . Finally, it should be noted that a continuously uniform, i.e. linear quantization by obtaining the above-described effect of deformation of the error PSD would also be possible, when the stage height would be adjusted so high or quantization so coarse that this quantization effectively works like a nonlinear quantization with regard to the signal statistic of the signal to be quantized, such as the prefiltered signal, wherein this stage height

adjustment is again made possible by the forward adaptivity of the prediction.

Further, the above-described embodiments can also be varied  
5 with regard to the processing of the encoded bit stream. Particularly, bit stream generator and extractor 214, respectively, could also be omitted.

The different quantizing indices, namely the residual  
10 values of the prefiltered signals, the residual values of the prefilter coefficients and the residual values of the prediction coefficients could also be transmitted in parallel to each other, stored or made available in another way for decoding, separately via individual channels. On  
15 the other hand, in the case that a constant bit rate is not imperative, these data could also be entropy-encoded.

Particularly, the above functions in the blocks of Figs. 1, 4, 5 and 6 could be implemented individually or in  
20 combination by sub-program routines. Alternatively, implementation of an inventive apparatus in the form of an integrated circuit is also possible, where these blocks are implemented, for example, as individual circuit parts of an ASIC.

25 Particularly, it should be noted that depending on the circumstances, the inventive scheme could also be implemented in software. The implementation can be made on a digital memory medium, particularly a disc or CD with  
30 electronically readable control signals, which can cooperate with a programmable computer system such that the respective method is performed. Generally, thus, the invention consists also in a computer program product having a program code stored on a machine-readable carrier  
35 for performing the inventive method when the computer program product runs on the computer. In other words, the invention can be realized as a computer program having a

program code for performing the method when the computer program runs on a computer.

Claims

1. An apparatus for encoding an information signal into an encoded information signal, comprising:
  - 5 a means (16) for determining a representation of a psycho-perceptibility motivated threshold, which indicates a portion of the information signal irrelevant with regard to perceptibility, by using a perceptual model;
  - 10 a means (18) for filtering the information signal for normalizing the information signal with regard to the psycho-perceptibility motivated threshold, for obtaining a prefiltered signal;
  - 15 a means (20) for predicting the prefiltered signal in a forward-adaptive manner to obtain a predicted signal, a prediction error for the prefiltered signal and a representation of prediction coefficients, based on which the prefiltered signal can be reconstructed; and
  - 20 a means (22) for quantizing the prediction error for obtaining a quantized prediction error, wherein the encoded information signal comprises information about the representation of the psycho-perceptibility motivated threshold, the representation of the prediction coefficients and the quantized prediction error.
  - 25
  - 30
2. The apparatus according to claim 1, wherein the means (22) for quantizing is implemented to quantize the prediction error via a quantizing function, which maps unquantized values of the prediction error to quantizing indices of quantizing stages, and whose course below a threshold is steeper than above a threshold.
- 35

3. The apparatus according to claim 1 or 2, wherein the means (22) for quantizing is implemented to obtain a quantizing stage height  $\Delta(n)$  of the quantizing function in a backward-adaptive manner from the quantized prediction error.
4. The apparatus according to one of the preceding claims, wherein the means (22) for quantizing the prediction error is implemented such that the unquantized values of the prediction error are quantized via clipping by the quantizing function, which maps the unquantized values of the prediction error to quantizing indices of a constant and limited first number of quantizing stages for obtaining the quantized prediction error.
5. The apparatus according to claim 4, wherein the means (22) for quantizing is implemented to obtain a quantizing stage height  $\Delta(n)$  of the quantizing function for quantizing a value  $(r(n))$  of the prediction error in a backward-adaptive manner of two past quantizing indices  $i_c(n-1)$  and  $i_c(n-2)$  of the quantized prediction error according to  $\Delta(n) = \beta \Delta(n-1) + \delta(n)$ , with  $\beta \in [0.0; 1.0]$ ,  $\delta(n) = \delta_0$  for  $|i_c(n-1) + i_c(n-2)| \leq I$  and  $\delta(n) = \delta_1$  for  $|i_c(n-1) + i_c(n-2)| > I$  with constant parameters  $\delta_0$ ,  $\delta_1$ ,  $I$ , wherein  $\Delta(n-1)$  represents a quantizing stage height obtained for quantizing a previous value of the prediction error.
6. The apparatus according to claims 4 or 5, wherein the means for quantizing is implemented to quantize the prediction error in a nonlinear manner.
7. The apparatus according to one of claims 4 to 6, wherein the constant and limited first number is 3.

8. The apparatus according to one of the preceding claims, wherein the means (16) for determining is implemented to determine the psycho-perceptibility motivated threshold in a block-wise manner from the information signal.  
5
9. The apparatus according to one of the preceding claims, wherein the means (16) for determining is implemented to represent the psycho-perceptibility motivated threshold in the LSF domain.  
10
10. The apparatus according to one of the preceding claims, wherein the means (16) for determining is implemented to determine the psycho-perceptibility motivated threshold in a block-wise manner and to represent the same in filtered coefficients, to subject the filter coefficients to a prediction and to subject a filter coefficient residual signal resulting from the prediction to a quantization via a further quantizing function, which maps the unquantized values of the filter coefficient residual signal to quantizing indices of quantizing stages, and whose course below a further threshold is steeper than above the further threshold, for obtaining a quantized filter coefficient residual signal, wherein the encoded information signal also includes information about the quantized filter coefficient residual signal.  
15  
20  
25
11. The apparatus according to claim 10, wherein means (16) for determining is implemented such that the unquantized values of the filter coefficient residual signal are quantized via clipping by the further quantizing function, which maps the unquantized values of the filter coefficient residual signal to quantizing indices of a constant and limited second number of quantizing stages.  
30  
35

12. The apparatus according to claim 11, wherein the means (16) for determining is implemented such that the prediction is performed in a backward-adaptive manner based on quantizing indices of the quantized filter coefficient residual signal.
13. The apparatus according to one of claims 10 to 12, wherein the means (16) for determining is implemented such that the prediction of the filter coefficients is performed by using a prediction filter with constant coefficients.
14. The apparatus according to one of claims 9 to 13, wherein the means (16) for determining is further implemented to subject the filter coefficients for representing the psycho-perceptibility motivated threshold to a subtraction with a constant value, prior to subjecting the same to prediction.
15. The apparatus according to one of the preceding claims, wherein the means (20) for predicting the prefiltered signal in a forward-adaptive manner further comprises:
- a means (36) for determining prediction filter coefficients from the prefiltered signal; and
- a means (44, 446, 48) for predicting the prefiltered signal via a filter (44) controlled by the prediction filter coefficients.
16. The apparatus according to claim 15, wherein the means (36) for determining is implemented to determine the prediction filter coefficients in a block-wise manner from the prefiltered signal.

17. The apparatus according to claim 15 or 16, wherein the means (36) for determining is implemented to represent the prediction filter coefficients in the LSF domain.
- 5 18. The apparatus according to one of claims 15 to 17, wherein the means (36) for determining is implemented to determine the prediction filter coefficients in a block-wise manner, to subject the prediction filter coefficients to a prediction, and to subject a  
10 prediction filter coefficient residual signal resulting from the prediction to quantization by a third quantizing function, which maps the unquantized values of the prediction filter coefficient residual signal to quantizing indices of quantizing stages, and  
15 whose course below a third threshold is steeper than above the third threshold, for obtaining a quantized prediction filter coefficient residual signal, wherein the encoded information signal also comprises information about the quantized prediction filter  
20 coefficient residual signal.
19. The apparatus according to claim 18, wherein means (36) for determining is implemented such that the unquantized values of the prediction filter  
25 coefficient residual signal are quantized via clipping to quantizing indices of the third number of quantizing stages by the third quantizing function, which maps the unquantized values of the prediction filter coefficient residual signal to quantize the  
30 indices of a constant and limited third number of quantizing stages.
20. The apparatus according to claim 18, wherein the means (36) for determining is implemented such that the  
35 prediction is performed in a backward-adaptive manner based on quantizing indices of the quantized prediction filter coefficients residual signal for one or several previous blocks of the prefiltered signal.

21. The apparatus according to one of claims 18 to 19,  
wherein the means (36) for determining is implemented  
such that the prediction of the prediction filter  
coefficients is performed by using a prediction filter  
with constant coefficients.
22. The apparatus according to one of claims 18 to 21,  
wherein the means (36) for determining is further  
implemented to subject the prediction filter  
coefficients to a subtraction with a constant value  
prior to subjecting the same to prediction.
23. The apparatus according to one of the preceding  
claims, which is implemented for encoding an audio  
signal or a video signal as information signal,  
wherein the perceptual model is a psychoacoustic model  
and the psycho-perceptibility motivated threshold a  
psychoacoustically motivated threshold, or the  
perceptual model is a psychovisual model and the  
psycho-perceptibility motivated threshold is a  
psychovisually motivated threshold.
24. An apparatus for decoding an encoded information  
signal comprising information about a representation  
of a psycho-perceptibility motivated threshold, a  
representation of prediction coefficients and a  
quantized prediction error into a decoded information  
signal, comprising:
- a means (206) for dequantizing the quantized  
prediction error for obtaining a dequantized  
prediction error;
- a means (208) for determining a predicted signal based  
on the prediction coefficients;

a means (210) for reconstructing a prefiltered signal based on the predicted signal and the dequantized prediction error; and

5 a means (212) for filtering the prefiltered signal for reconverting a normalization with regard to the psycho-perceptibility motivated threshold for obtaining the decoded information signal.

10 25. The apparatus according to claim 24, wherein the means (206) for dequantizing is implemented to dequantize the quantized prediction error to a limited and constant number of quantizing stages.

15 26. The apparatus according to claim 25, wherein the means (206) for dequantizing is implemented to obtain a quantizing stage height  $\Delta(n)$  between the quantizing stages in a backward-adaptive manner from already dequantized quantizing indices of the quantized prediction error.  
20

27. The apparatus according to claim 25 or 26, wherein the means (260) for dequantizing is implemented to obtain a quantizing stage height ( $\Delta(n)$ ) between the  
25 quantizing stages for dequantizing a quantizing index of the quantized prediction error in a backward-adaptive manner from two previous quantizing indices  $i_c(n-1)$  and  $i_c(n-2)$  of the quantized prediction error according to  $\Delta(n) = \beta\Delta(n-1) + \delta(n)$  with  
30  $\beta \in [0.0; 1.0]$ ,  $\delta(n) = \delta_0$  for  $|i_c(n-1) + i_c(n-2)| \leq I$  and  $\delta(n) = \delta_1$  for  $|i_c(n-1) + i_c(n-2)| > I$  having constant parameters  $\delta_0$ ,  $\delta_1$ ,  $I$ , wherein  $\Delta(n-1)$  represents a quantizing stage height obtained for dequantizing  
35  $i_c(n-1)$ .

28. The apparatus according to one of claims 25 to 27, wherein the constant and limited number is less than or equal to 32.

29. The apparatus according to one of claims 25 to 28,  
wherein the constant and limited number is 3.
- 5 30. The apparatus according to one of claims 24 to 29,  
wherein the means (212) for filtering comprises:
- 10 a means (230) for determining perceptual threshold  
filter coefficients from the information about the  
representation of the psycho-perceptibility motivated  
threshold in a block-wise manner for blocks of a  
sequence of blocks of the prefiltered signal; and
- 15 a postfilter (232) for filtering the prefiltered  
signal by using the perceptual threshold filter  
coefficients.
- 20 31. The apparatus according to one of claims 24 to 30,  
wherein the means (230) for determining is implemented  
to obtain the perceptual threshold filter coefficients  
by reversion from an LSF domain.
- 25 32. The apparatus according to one of claims 24 to 31,  
wherein the means (230) for determining is implemented  
to obtain quantizing indices of a quantized filter  
coefficient residual signal from the representation of  
the psycho-perceptibility motivated threshold, to  
dequantize the same to a limited and constant second  
number of quantizing levels, for obtaining a  
30 dequantized filter coefficient residual signal, to  
predict the filter coefficients representing the  
psycho-perceptibility motivated threshold and to add  
the same to the dequantized filter coefficient  
residual signal and to convert a reconstructed filter  
35 coefficient residual signal resulting from the  
addition by reversion into the perceptual threshold  
filter coefficients.

33. The apparatus according to claim 32, wherein the means (230) for determining is implemented such that the prediction is performed in a backward-adaptive manner based on already predicted filter coefficients representing the psycho-perceptibility motivated threshold.
34. The apparatus according to claims 32 or 33, wherein the means (230) for determining is implemented such that the prediction of the filter coefficients representing the psycho-perceptibility motivated threshold is performed by using a prediction filter with constant coefficients.
35. The apparatus according to one of claims 32 to 34, wherein the means (230) for determining is further implemented to subject the reconstructed filter coefficient residual signal resulting from the addition to an addition with a constant value prior to reconversion.
36. The apparatus according to one of claims 24 to 37, wherein the means (208) for determining a predicted signal further comprises:
- a means (224) for determining prediction filter coefficients from the representation of prediction coefficients comprised in the encoded information signal; and
- a means (226, 228) for predicting the prefiltered signal via a filter (226) controlled by the prediction filter coefficients.
37. The apparatus according to claim 36, wherein the means (224) for determining prediction filter coefficients is implemented to determine the same in a block-wise

manner for blocks of a sequence of blocks of the prefiltered signal.

38. The apparatus according to one of claims 36 or 37,  
5 wherein the means (224) for determining is implemented to obtain the prediction filter coefficients by reconversion from an LSF domain.
39. The apparatus according to one of claims 36 to 38,  
10 wherein the means (224) for determining is implemented to obtain quantizing indices of a quantized prediction coefficient residual signal from the representation of the prediction coefficients, to dequantize the same to  
15 levels for obtaining a dequantized prediction coefficient residual signal, to predict prediction filter coefficients and to add the same to the dequantized prediction coefficient residual signal and to convert a reconstructed prediction coefficient  
20 residual signal resulting from the addition by reconversion into the prediction filter coefficients.
40. The apparatus according to claim 39, wherein the means  
25 (224) for determining is implemented such that the prediction is performed in a backward-adaptive manner based on the already predicted prediction coefficients.
41. The apparatus according to claim 39 or 40, wherein the  
30 means (224) for determining is implemented such that the prediction of the prediction coefficients is performed by using a prediction filter with constant coefficients.
- 35 42. The apparatus according to one of claims 39 to 41, wherein the means (224) for determining is further implemented to subject the reconstructed prediction coefficient residual signal resulting from the

addition to an addition with the constant value prior to reconversion.

5 43. The apparatus according to one of claims 24 to 42, which is implemented for decoding an audio signal or a video signal as information signal, and wherein the psycho-perceptibility motivated threshold is an acoustic masking threshold or a visual masking threshold.

10

44. A method for encoding an information signal into an encoded information signal, comprising:

15 using a perceptibility model, determining a representation of a psycho-perceptibility motivated threshold indicating a portion of the information signal irrelevant with regard to perceptibility;

20 filtering the information signal for normalizing the information signal with regard to the psycho-perceptibility motivated threshold for obtaining a prefiltered signal;

25 predicting the prefiltered signal in a forward-adaptive manner to obtain a prefiltered signal, a prediction error to the prefiltered signal and a representation of prediction coefficients, based on which the prefiltered signal can be reconstructed; and

30 quantizing the prediction error to obtain a quantized prediction error, wherein the encoded information signal comprises information about the representation of the psycho-perceptibility motivated threshold, the representation of the prediction coefficients and the  
35 quantized prediction error.

45. A method for decoding an encoded information signal comprising information about the representation of a

psycho-perceptibility motivated threshold, a representation of prediction coefficients and a quantized prediction error into a decoded information signal, comprising:

5

dequantizing the quantized prediction error to obtain a dequantized prediction error;

10

determining a predicted signal based on the prediction coefficient;

15

reconstructing a prefiltered signal based on the predicted signal and the dequantized prediction error; and

20

filtering the prefiltered signal for converting a normalization with regard to the psycho-perceptibility motivated threshold to obtain the decoded information signal.

46. A computer program with a program code for performing the method according to claim 44 or 45 when the computer program runs on a computer.

25 47. An encoder, comprising:

an information signal input (12);

30

a perceptibility threshold determiner (26) operating according to a probability model having an input coupled to the information signal input and a perceptibility threshold output;

35

an adaptive prefilter (34) comprising a filter input coupled to the information signal input, a filter output and a adaption control input coupled to the perceptibility threshold output,

- a forward prediction coefficient determiner (36) comprising an input coupled to the prefilter output and a prediction coefficient output;
- 5 a first differentiator (42) comprising a first differentiator input coupled to the prefilter output, a second differentiator input and a differentiator output;
- 10 a clipping/quantizing stage (52) comprising a limited and constant number of quantizing levels, an input coupled to the differentiator output, a quantizing step size control input and an output;
- 15 a step size adjuster (54) comprising an input coupled to the output of the clipping/quantizing stage (52) and a quantizing step size output coupled to the quantizing step size control input of the clipping/quantizing stage (52);
- 20 a dequantizing stage (50) comprising an input coupled to the output of the clipping/quantizing stage and a dequantizer control output;
- 25 an adder (48) comprising a first adder input coupled to the dequantizer output, a second adder input and an adder output;
- 30 a prediction filter (44, 46) comprising a prediction filter input coupled to the adder output, a prediction filter output coupled to the second differentiator input as well as to the second adder input, as well as a prediction coefficient input coupled to the prediction coefficient output;
- 35 an information signal generator (24) comprising a first input coupled to the perceptibility threshold output, a second input coupled to the prediction

coefficient output, a third input coupled to the output of the clipping/quantizing stage and an output representing an encoder output.

- 5    48. A decoder for decoding an encoded information signal comprising information about a representation of a psycho-perceptibility motivated threshold, prediction coefficients and a quantized prediction error, into a decoded information signal, comprising:
- 10        a decoder input;
- an extractor (214) comprising an input coupled to the decoder input, a perceptibility threshold output, a
- 15        prediction coefficient output and a quantized prediction error output;
- a dequantizer (206) comprising a limited and constant number of quantizing levels, a dequantizer input
- 20        coupled to the quantized prediction error output, a dequantizer output and a quantizing threshold control input;
- a backward-adaptive threshold adjuster comprising an
- 25        input coupled to the quantized prediction error output, and an output coupled to the quantized threshold control input;
- an adder (222) comprising a first adder input coupled
- 30        to the dequantizer output, a second adder input and an adder output;
- a prediction filter (226) comprising a precision filter input coupled to the adder output, a prediction
- 35        filter output coupled to the second input, and a prediction filter coefficient input coupled to the prediction coefficient output; and

an adaptive postfilter (232) comprising a prediction filter input coupled to the adder output, a prediction filter output representing a decoder output, and an adaption control input coupled to the perceptibility threshold output.

5