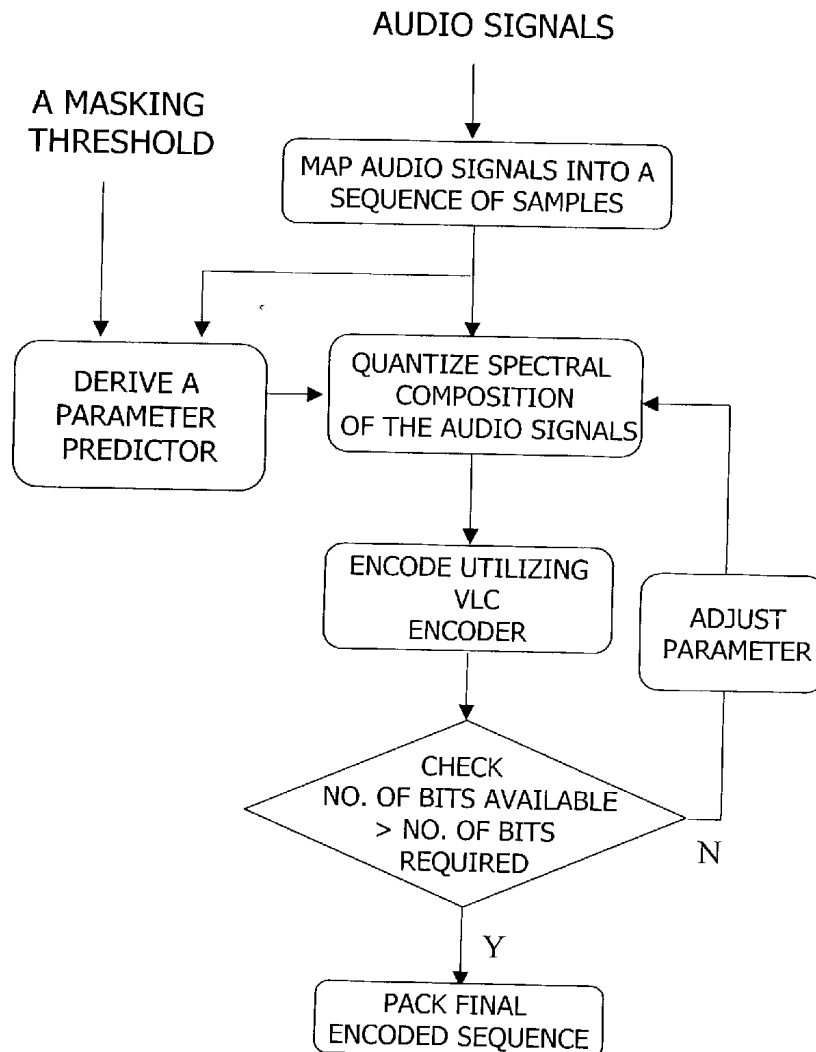




US 20040002859A1

(19) **United States**(12) **Patent Application Publication****Liu et al.**(10) **Pub. No.: US 2004/0002859 A1**(43) **Pub. Date: Jan. 1, 2004**(54) **METHOD AND ARCHITECTURE OF
DIGITAL CODING FOR TRANSMITTING
AND PACKING AUDIO SIGNALS**(76) Inventors: **Chi-Min Liu, Hsinchu Hsien (TW);
Wen-Chieh Lee, Taoyuan Hsien (TW)**Correspondence Address:
**SUPREME PATENT SERVICES
POST OFFICE BOX 2339
SARATOGA, CA 95070 (US)**(21) Appl. No.: **10/184,157**(22) Filed: **Jun. 26, 2002****Publication Classification**(51) **Int. Cl.⁷ G10L 19/02**(52) **U.S. Cl. 704/229**(57) **ABSTRACT**

A method of digital coding transforms input audio signals into a sequence of frequency samples representing a spectral composition of the audio signals, and quantizes the sequence of frequency samples into quantized values according to a bit allocation process which uses a parameter predictor to evaluate quantization parameters by referring to a masking threshold. The quantized values are encoded into a number of bits of encoded data. An iterative rate control loop adjusts the quantization parameters and the quantization step size if the number of bits in the encoded data exceeds a prescribed number of bits available for the encoded data. The method may also cut off high frequency components of the input audio signals according to a cut-off frequency determined by the iterative rate control loop before quantizing the sequence of frequency samples.



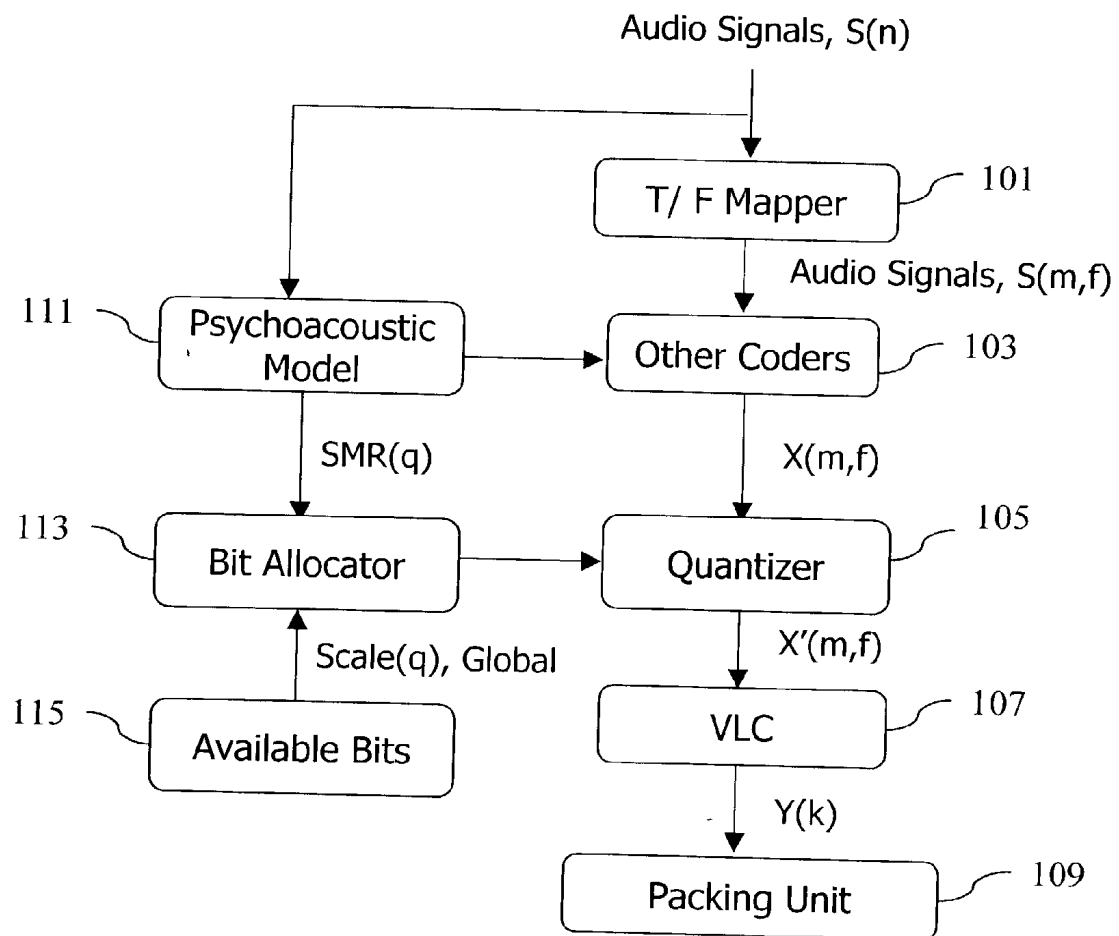


FIG. 1 (Prior Art)

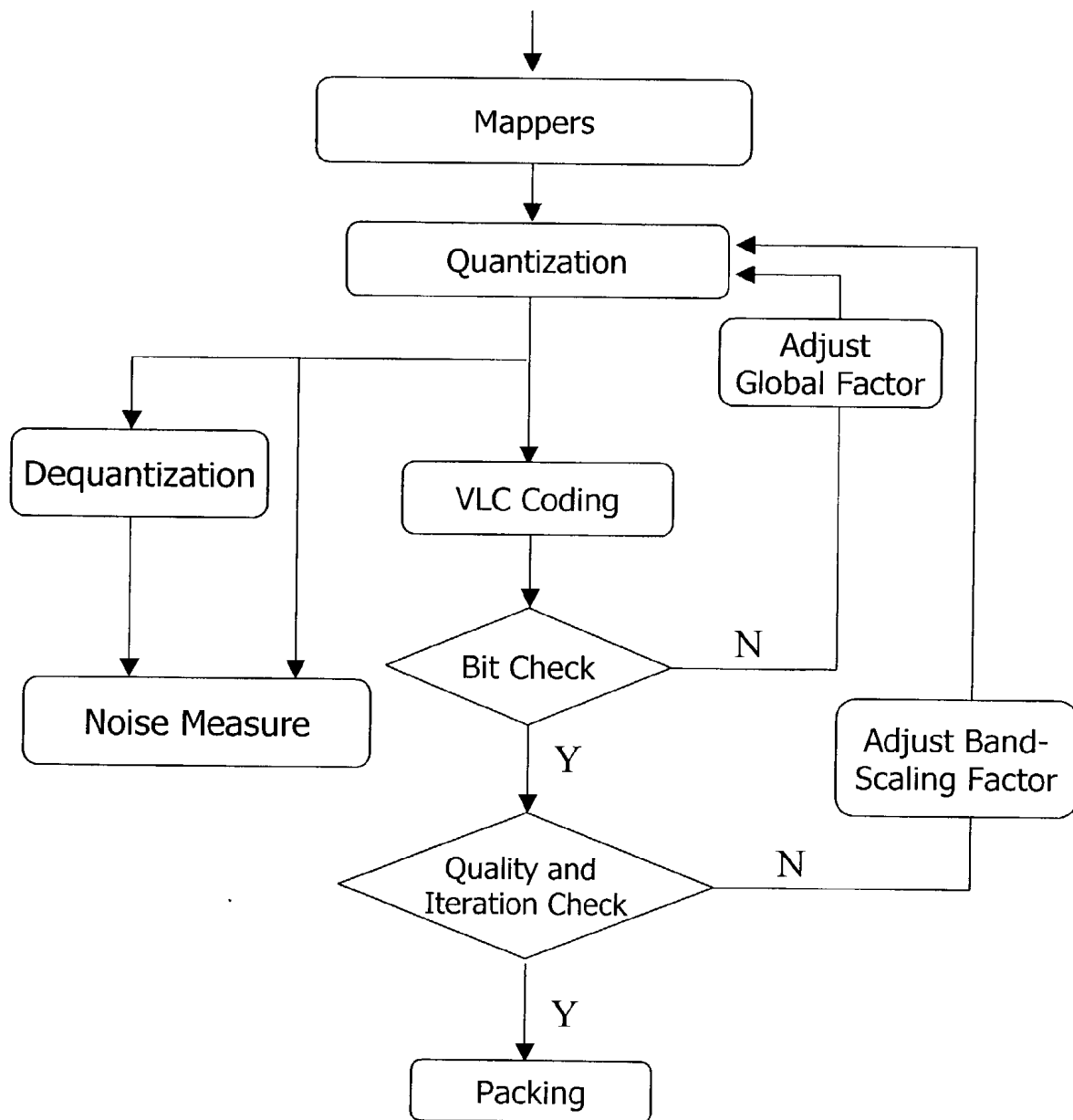


FIG. 2 (Prior Art)

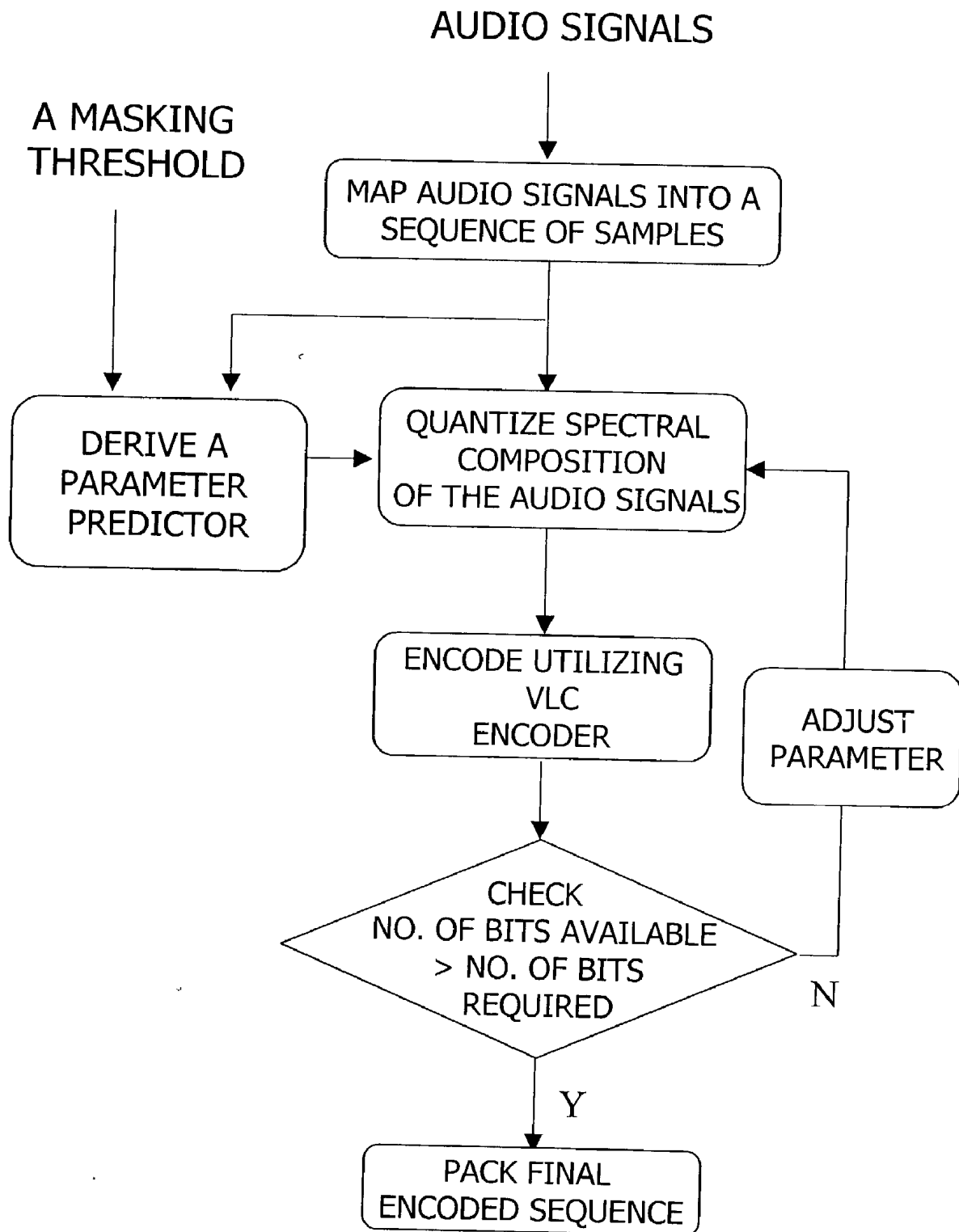


FIG. 3a

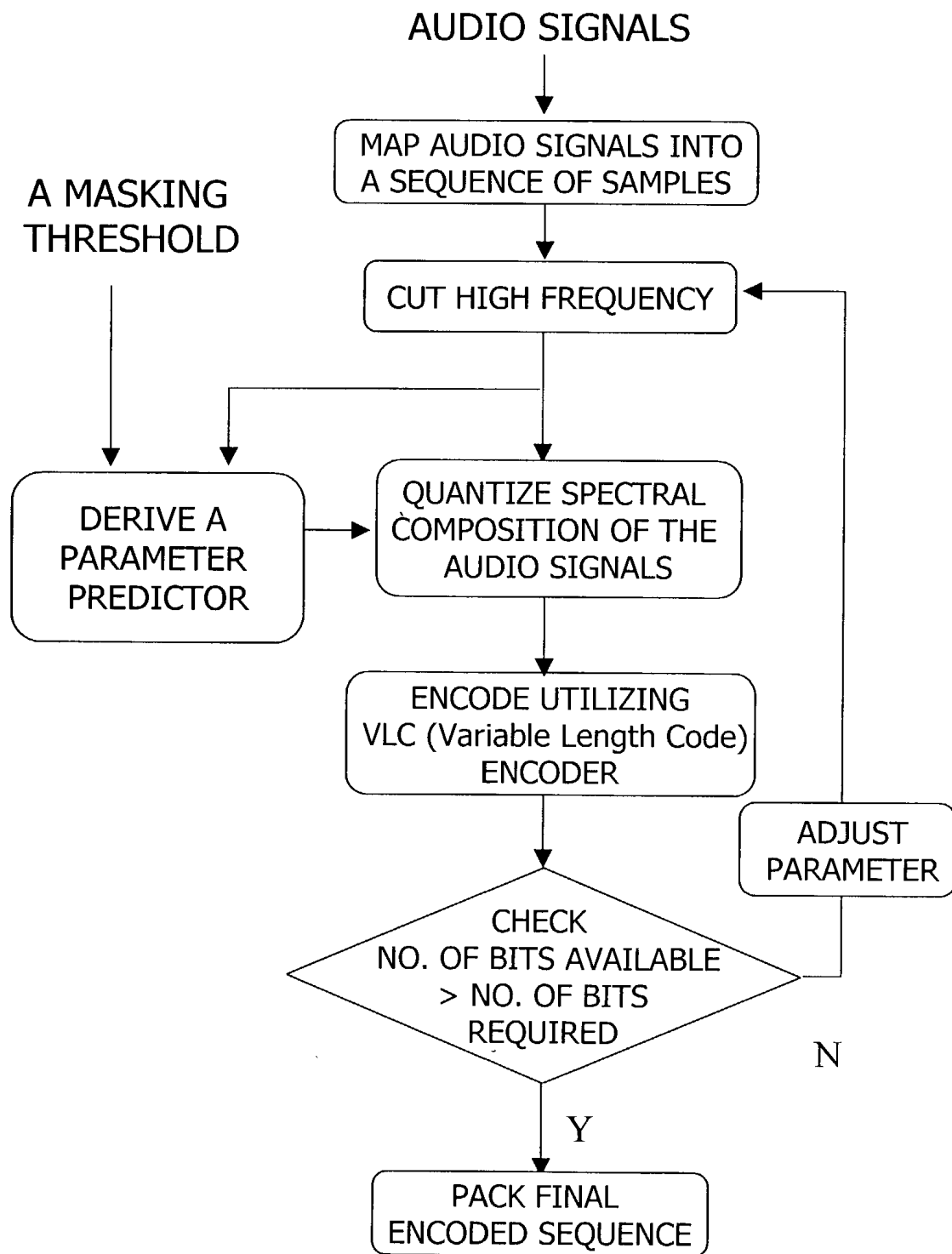


FIG. 3b

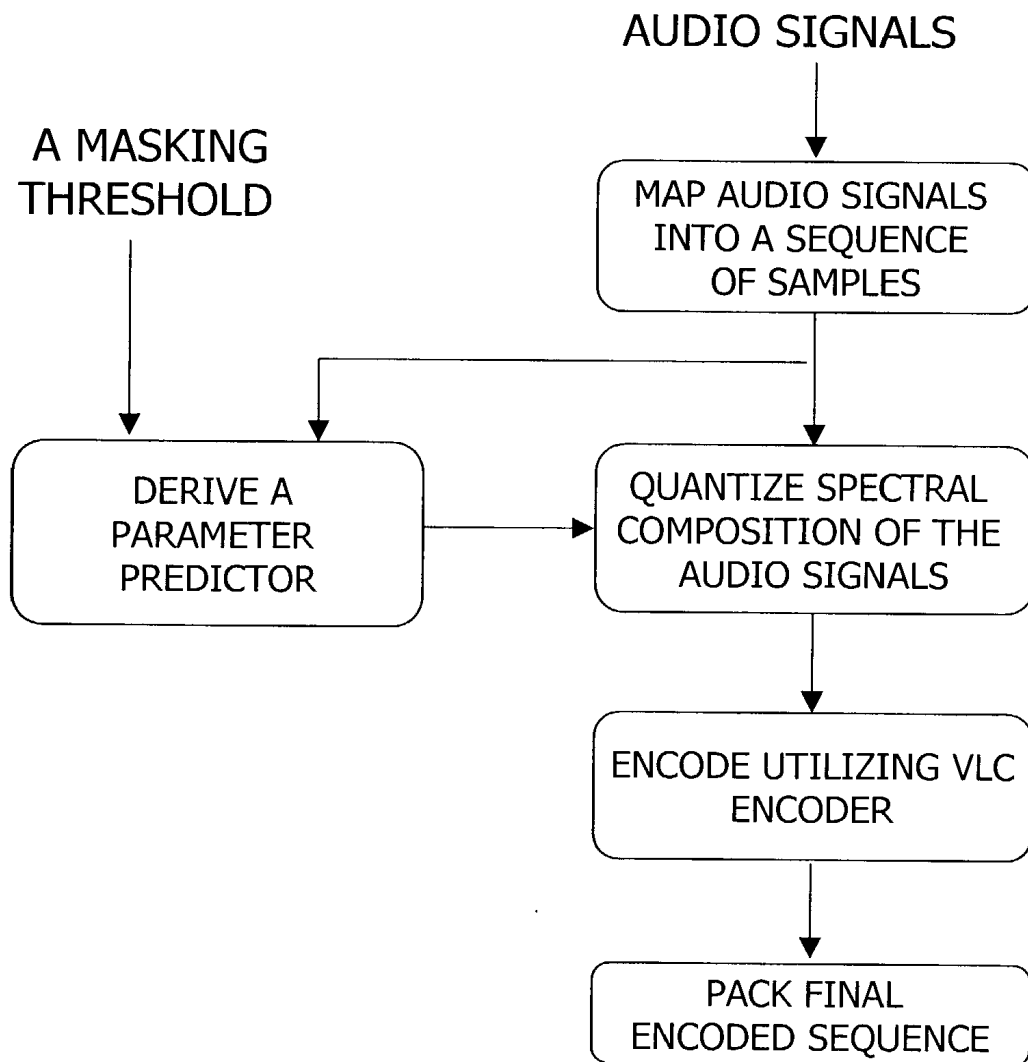


FIG. 3c

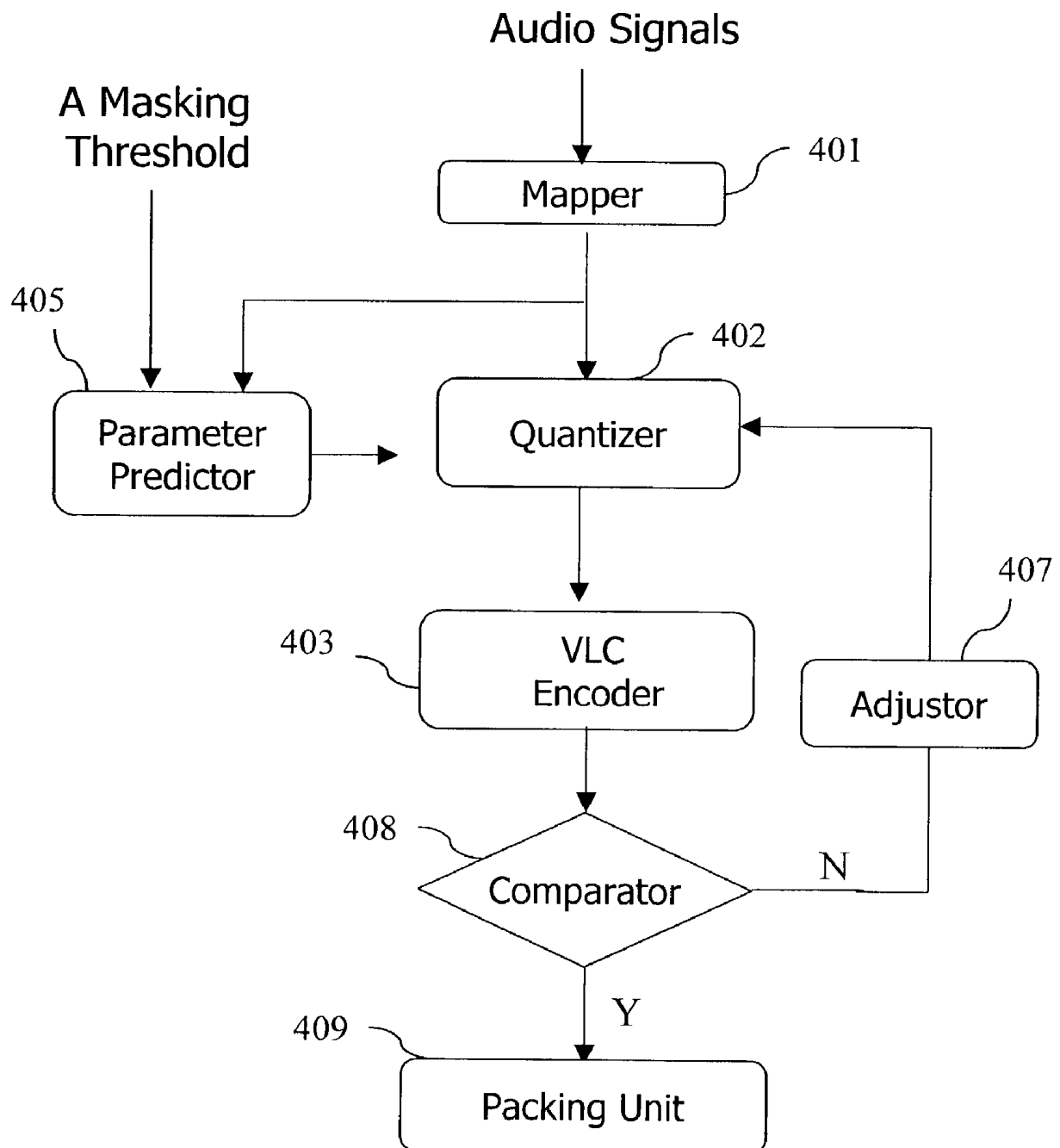


FIG. 4a

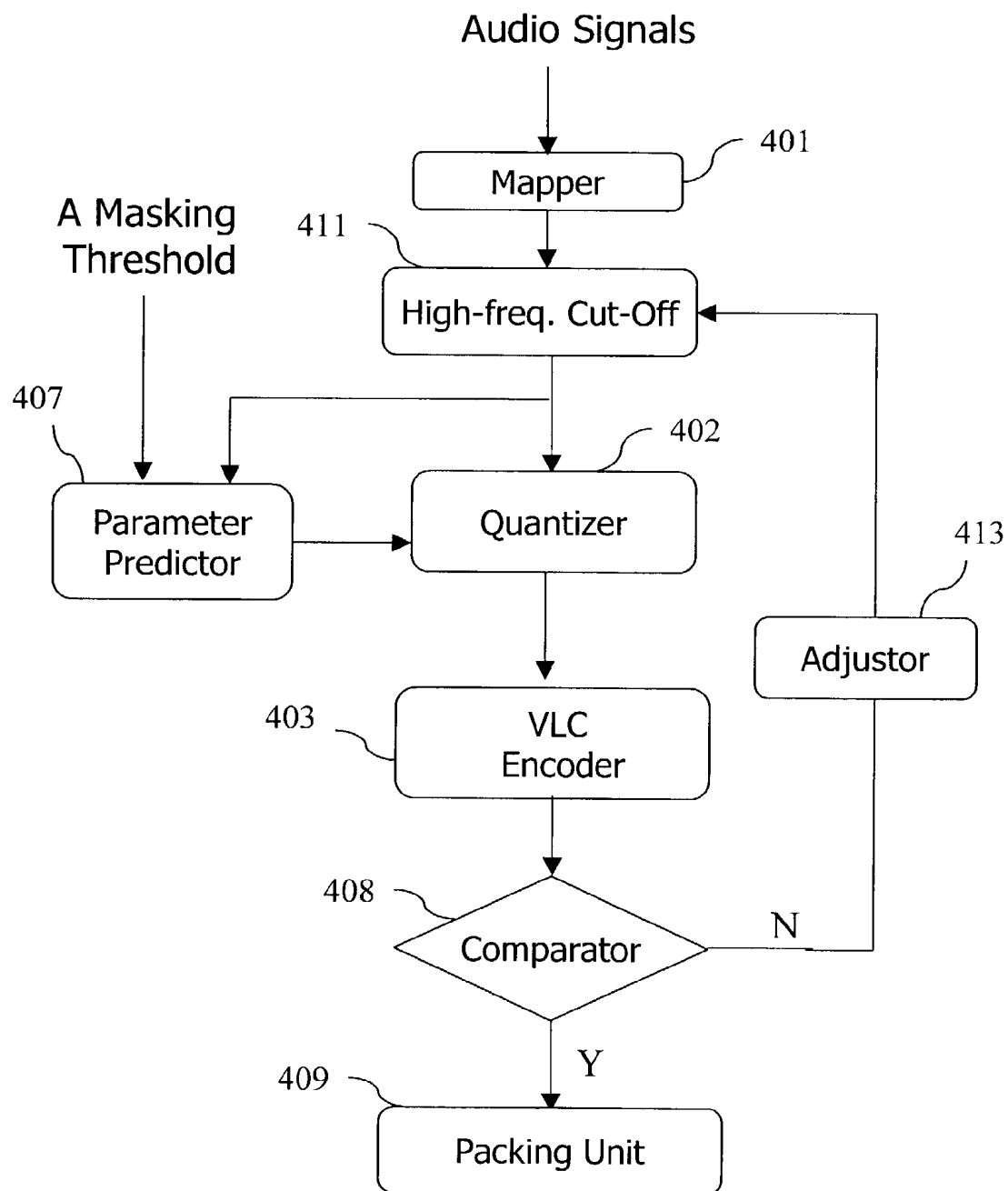


FIG. 4b

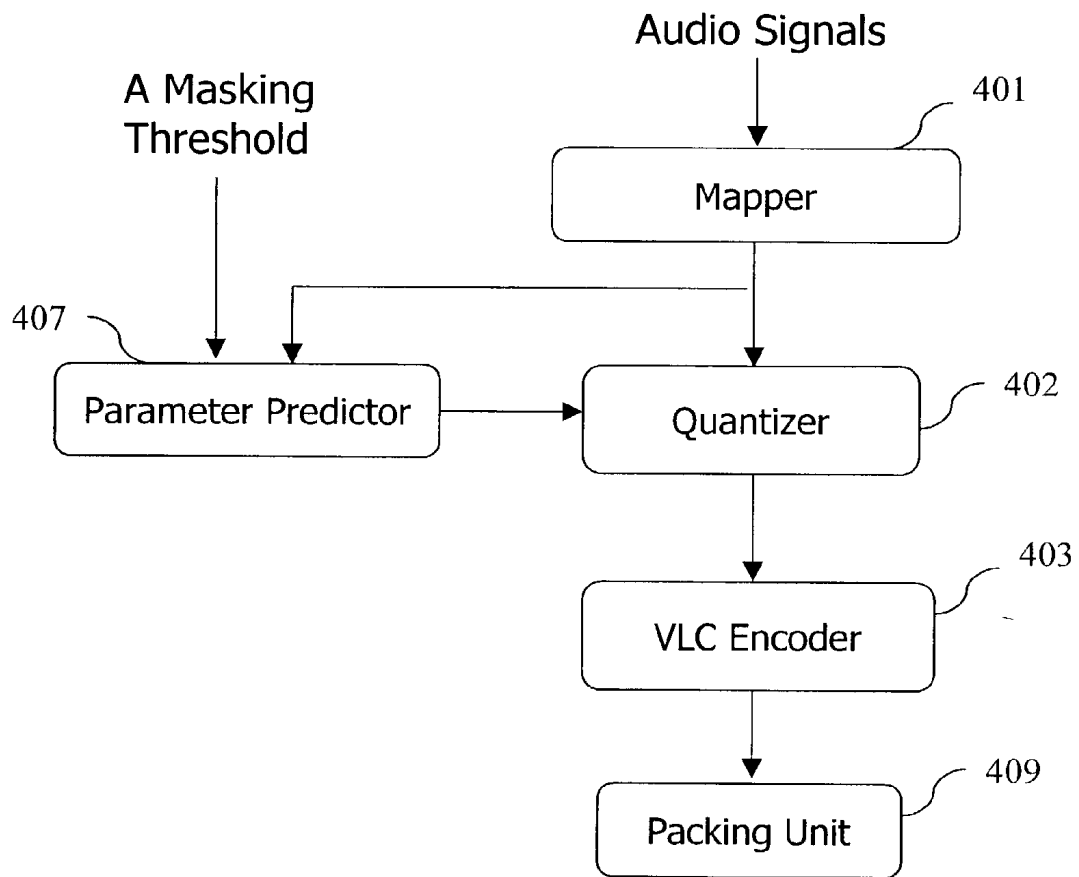


FIG. 4c

Track	1	2	3	4	5	6	7	8
R in ISO(128kbps)	38.33	39.52	29.66	40.77	37.79	39.39	38.98	38.07
Q in ISO(128kbps)	11.39	13.76	8.97	11.32	11.10	11.71	12.68	11.63
R in the invention	12.20	12.58	11.98	11.33	12.99	12.90	11.61	11.30
Q in the invention	0	0	0	0	0	0	0	0

FIG. 5

	1	2	3	4	5	6	7	8
ISO	-2.11	-2.32	-2.06	-1.48	-2.42	-2.22	-1.28	-1.67
ISO1	-1.84	-1.88	-1.58	-1.19	-2.44	-1.88	-0.99	-1.27
present invention	-1.71	-1.35	-1.87	-0.82	-1.93	-1.64	-0.94	-1.17

FIG. 6

No.	Item	File name	Remarks
1	Bass	Bass	4
2	Glockenspiel	Gspi	1 & 2 & 5
3	Harpsichord	Hach	1 & 2 & 4
4	Horn	Horn	4
5	Quarter	Quar	4
6	Soprano	Sopr	4
7	Trumpet	Trpt	2
8	Violoncello	Vioo	5

Remarks:

- 1) Transients: pre-echo sensitive, smearing of noise in temporal domain.
- 2) Tonal structure: noise sensitive, roughness.
- 3) Natural speech (critical combination of tonal parts and attacks): distortion sensitive, smearing of attacks.
- 4) Complex sound: stresses the Device Under Test.
- 5) High bandwidth: stresses the Device Under Test, loss of high frequencies, programme-modulated high frequency noise.

FIG. 7

METHOD AND ARCHITECTURE OF DIGITAL CODING FOR TRANSMITTING AND PACKING AUDIO SIGNALS

FIELD OF THE INVENTION

[0001] The present invention relates generally to a method and its architecture of digital coding for transmitting and packing signals and, in particular, to the bit allocation in the coding of audio signals.

BACKGROUND OF THE INVENTION

[0002] The perceptual audio coding such as MPEG Layers 1-3, advanced audio coding, or T/F (Time/Frequency) coding, has been widely used in consumer electronics, telecommunications, and broadcasting. Among these perceptual audio coders, the bit allocation is one of the main tasks leading to the high complexity and the key module determining encoded quality.

[0003] FIG. 1 illustrates the block diagram of a coding process in perceptual audio coding. A T/F mapper 101 transforms the audio signals $S(n)$ into frequency segments $S(m, f)$ from time domain into frequency domain by a window-by-window basis. Various coders 103 have been used in the coding process to achieve high compression ratios. The output $X(m, f)$ is the frequency domain sequence after coding with the window segment index m and the frequency index f . A quantizer 105 quantizes $X(m, f)$ into a finite number of levels represented by $X'(m, f)$ with the goal of minimizing the subjective impairments introduced by the quantization noise. The quantization levels are controlled through the quantization parameters.

[0004] The audio compression in general classifies the frequency lines into sets referred to as quantization bands. The number of lines grouped in a quantization band is determined according to the critical bands and the affordable bits that are required to transmit the quantization parameters. VLC (Variable length coding) 107 represents the quantized sequence $X'(m, f)$ through a variable length coding with the consideration of the statistic occurrence probability of the transmitted signal. A packing unit 109 packs the final encoded sequence into a sequence defined by a specified audio protocol. A psychoacoustic model 111 analyzes the signals and provides SMR (signal-to-masking ratio) for the quantization bands from the signal analysis result. A bit-allocator 113 determines the quantization parameters with reference to the masking thresholds provided by the psychoacoustic model 111 and the available bit budget 115.

[0005] A non-uniform quantizer quantizes the spectral lines under the control of the bit allocator, which decides the quantization manners with the consideration of the resultant audio quality and the required bits. Hence control over the quality and the bit number is the fundamental requirement of the bit allocation. U.S. Pat. No. 5,579,430 discloses a digital encoding process related to the OCF (optimum coding in the frequency domain) process. It improves the OCF process in such a manner that encoding of music with a quality comparable to compact-disc quality is possible at a data rate of approximately 2 bits/ATW and with good FM-radio quality at a data rates of 1.5 bits/ATW. Another U.S. Pat. No. 5,924,060 discloses a digital coding process for the transmission and/or storage of acoustical signals, which reduces

the data rate by a factor of 4 to 6 without subjectively degrading the quality of the musical signal.

[0006] For MPEG Layers 1 and 2, a uniform quantizer is utilized to control the quality and the bit requirement. Hence the bit allocation is simply to apportion the total number of bits available for the quantization of the sub-band signals to minimize the audibility of the quantization noise. For coders such as MPEG Layer 3, MPEG-2 AAC, and MPEG4 T/F coding, control over the quality and the bit rate is difficult. This is mainly due to the fact that they all use a non-uniform quantizer whose quantization noise varies with respect to the input values. In other words, it fails to control the quality by assigning quantizer parameters according to the perceptually allowable noise. In addition, the variable length coding used in MPEG Layer 3 and MPEG-2 AAC assigns variable bit-length to different values, which means that the bits consumed should be obtained from the quantization results, and cannot be from the quantizer parameters alone. Thus, the bit allocation is one of the main tasks leading to the high complexity of the encoder.

[0007] The above drawbacks lead to the problem in evaluating the quantization parameters. A two-nested loop iterative method referred to as the OCF has been proposed to solve the problem. As illustrated in FIG. 2, it evaluates the quantization parameters through two iteration loops, the rate-controlling loop and the quality-controlling loop. The rate-controlling loop iteratively adjusts the parameter values to fit to the limited bits obtained by performing quantization and Huffman coding for spectral lines. The quality-controlling loop iteratively adjusts the parameter values to fit to a perceptual criterion of the quantization noise that needs to be evaluated by performing the inverse quantization.

[0008] The complexity of the method for a frame with F spectral lines can be described as $O(F \cdot R \cdot \eta + F \cdot Q \cdot \gamma)$, where Q and R are respectively the numbers of quality-controlling iterations and rate-controlling iterations while η and γ are the computation complexity to handle a spectral line in the rate-controlling loop and the quality-controlling loop, respectively. The rate-controlling loop complexity η is from the quantization and the VLC coding of a spectral line while the quality-controlling loop complexity γ is from the dequantization and noise measure. Both complexity η and γ are high. Also, the numbers of iterations Q and R depend on the initial values of quantization parameters and the adjustment methods. The complexity is even larger than the total complexity of the hybrid transform and the psychoacoustic model shown in FIG. 1.

[0009] Assigning bits to quantization bands in the quality-controlling loop determines the quality of the coded audio. There have been two approaches to assigning the bits. One approach is to assign the bit only to the band with the worst noise-to-masking ratio in each of the iterations in the loop. The approach leads to a large number of iterations in the quality-controlling loop, which means very high complexity. Another approach assigns bits to all the bands with a noise-to-masking ratio higher than one in each of the iteration until all available bits are consumed. This approach has a much lower complexity than the first approach. However, whether the quality of the approach is satisfactory is the concerns.

[0010] The first approach can shape the noise so that the masking threshold will be in parallel to the noise threshold,

which has been a widely accepted criterion. The second approach that has been in the sample code provided by ISO usually leads to better subjective quality. The problems of the two nested loops method is that it may not lead to a convergent condition. Since there are two separate rules controlling the quality and bits consumed in two loops, it may lead to infinite loops, generally referred to as dead-lock problem. A general method to manage the deadlock problem is to set a limit to the maximum number of iterations, and use some heuristic parameter tuning method to take care of the quality and the loop number. However, the quality can not be guaranteed for these methods.

SUMMARY OF THE INVENTION

[0011] This invention has been made to overcome the drawbacks of the conventional digital coding process. The primary object is to provide a method of digital coding for transmitting and packing audio signals with high quality and much less computing complexity.

[0012] According to the invention, input audio signals are first mapped into a sequence of frequency samples to represent a spectral composition of the audio signals. The sequence of frequency samples is quantized in accordance with a bit allocation process and a parameter predictor evaluating the quantization parameters by directly referring to a masking threshold. These quantized values are encoded with variable length coding or directly packed to a specified protocol. If the overall length of the encoded data exceeds the number of bits available, a parameter adjustment is made and the quantization step size is increased. This process is repeated until the number of bits available is greater than the number of required bits for the encoding. Finally, the final encoded sequence is packed into a sequence defined by a specified audio protocol.

[0013] The method of this invention takes a non-uniform quantizer of MPEG layer 3 for detail derivation and examines the issues of the complexity and audio quality of the perceptual encoding method. Accordingly, it uses segmental-noise-to-masking-ratio for the derivation, and provides a closed-form equation for the relationship between bits/step size and quantization noise. The method is not limited to MPEG Layer 3, it is applicable to most perceptual coders like MPEG AAC (advanced audio coding). It is also applicable to the coder with uniform quantizers such as MPEG Layer 1 and Layer 2 due to the new bit allocation criteria this invention provides.

[0014] Another object of the present invention is to provide the architecture for such a digital coding process. The architecture comprises a mapper, a quantizer, a VLC encoder, a parameter predictor, a packing unit, an adjustor, and a comparator that may be realized by signal processors to accomplish the method of this invention.

[0015] According to the present invention, the quantization parameters are evaluated directly from the quality criteria for the graceful degradation in consideration of the quantization bandwidth and the required bits in the non-equal frequency lines by means of a rate-controlling loop for low bit-rate audio coding process. For variable bit-rate coding, the iteration in rate-controlling loop can be removed completely.

[0016] The foregoing and other objects, features, aspects and advantages of the present invention will become better

understood from a careful reading of a detailed description provided herein below with appropriate reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0017] **FIG. 1** illustrates the block diagram of a coding process in modern audio coding.

[0018] **FIG. 2** illustrates the bit allocation process for an OCF process.

[0019] **FIG. 3a** illustrates the procedure of the audio coding process according to the present invention.

[0020] **FIG. 3b** illustrates the procedure of the low bit-rate audio coding process according to the present invention.

[0021] **FIG. 3c** illustrates the procedure of the variable bit-rate audio coding process according to the present invention.

[0022] **FIG. 4a** illustrates a realized architecture of **FIG. 3a** according to the present invention.

[0023] **FIGS. 4b** and **4c** illustrate the realized architectures of **FIGS. 3b** and **3c** respectively.

[0024] **FIG. 5** illustrates the average iteration number for each granule in MPEG Layer 3 with different testing material for the present invention and the MPEG bit allocation process respectively.

[0025] **FIG. 6** illustrates the objective score of the method of the invention compared to the bit allocation method suggested in ISO draft.

[0026] **FIG. 7** provides a list with a subset of test signals that were used during the objective and subjective test.

DETAILED DESCRIPTION OF THE INVENTION

[0027] **FIG. 3a** illustrates the procedure of the audio coding method according to the present invention. Referring to **FIG. 3a**, input audio signals are first mapped into a sequence of frequency samples representing a spectral composition of the audio signals. This sequence of frequency samples is then quantized to obtain symbols with a lower precision according to a bit allocation process. A parameter predictor is used to evaluate the quantization parameters by directly referring to a masking threshold for the noise extent that a human hearing system can hear. The parameters determining the signal level resolution for a compression system are predicted.

[0028] These quantized symbols are encoded with a VLC encoder. The next step is checking if a prescribed number of bits available is enough or not for the encoded data. If the number of bits available is not greater than the overall length of the encoded data, a parameter adjustment is made and the quantization step size is increased. This process is repeated until the number of required bits for the encoding reaches the number of bits available. At the end, the final encoded sequence is packed into a sequence defined by a specified audio protocol.

[0029] For audio coding of a low bit-rate, the high frequency may be cut off before evaluating the quantization parameters in the parameter predictor. **FIG. 3b** illustrates the procedure of the low bit-rate audio coding process. As

shown in **FIG. 3b**, while the number of required bits for the low bit-rate encoding exceeds the number of bits available, the cut-off frequency is adjusted and transmitted so that the high frequency components are cut off before evaluating the quantization parameters. The quantization step size may also be adjusted if desirable. For audio coding of a variable bit-rate, the available bits can be adjusted according to the required quality. In this case, the iteration in the rate control loop can be completely removed. **FIG. 3c** illustrates the procedure of the variable bit-rate audio coding process, in which the iteration in the rate control loop is removed from **FIG. 3a**.

[0030] The procedures as shown in **FIGS. 3a-3c** of this invention may be realized with signal processors. The detailed architectures of the realization are disclosed as follows. In accordance with **FIG. 3a**, the realized architecture shown in **FIG. 4a** comprises a mapper **401** to receive and transform an input sequence of audio signals into a sequence of frequency samples to thereby represent a spectral composition of the audio signals. A quantizer **402** quantizes the sequence of frequency samples into a finite number of levels in accordance with a bit allocation process. A parameter predictor **405** is used to evaluate the quantization parameters by directly referring to a masking threshold, and an optimum encoder **403** encodes the quantized levels. An adjustor **407** adjusts the quantization parameters when the number of bits available is not enough for the encoded data and a comparator **408** compares a prescribed number of bits available and the required length of the encoded data to check if the number of bits available is enough or not for the encoded data. A packing unit **409** packs the final encoded sequence into a sequence defined by a specified audio protocol.

[0031] **FIGS. 4b** and **4c** illustrate the realized architectures of **FIGS. 3b** and **3c** respectively. Referring to **FIG. 4b**, an adjustor **413** is used to adjust the cut-off frequency and transmit it to a high-frequency cut-off unit **411** in the case of low bit-rate audio coding. The adjustor **413** may also adjust the quantization step size used in the quantizer **402**. The high-frequency cut-off unit **411** is added between the mapper **401** and the quantizer **402** to receive the adjusted cut-off frequency and transmit it to the parameter predictor **405**. In the case of variable bit-rate coding, the elements related to the iteration in the rate control loop are simply removed as shown in **FIG. 4c**.

[0032] In the invention, a deterministic formula based on a constant masking-to-noise ratio ρ is derived to calculate the quantization parameters for the parameter predictor in the bit allocation process. It provides a closed-form equation of the noise predictor for a non-uniform quantizer. This invention takes MPEG Layer 3 as the detailed derivation and experiment example. For a MPEG ACC quantizer, a similar process is applicable.

[0033] The bit allocation of the present invention meets the requirement of bit rate and noise shaping for each sub-band by single step prediction. An optimum global factor and a scaling factor for each sub-band are evaluated by directly referring to a masking threshold. The global factor controls the overall number of consumed bits, and the scaling factor controls the quantization noise of the associated band relative with the other bands. The following paragraphs first illustrate the bit allocation criteria, then

derive in more detail the noise predictor and bounds on a scale factor under the constraint from the zero band and negative noise-to-masking ratio (NMR).

[0034] Bit Allocation Criteria

[0035] Firstly, the minimum over segmental NMR is considered:

$$R(i) = \arg \min_{R(i)} \sum_i \left\{ \left(\frac{\sigma_{N(i)}^2}{\sigma_{M(i)}^2} \right) \right\}, \quad (1)$$

[0036] where

$$\sigma_{N(i)}^2 \text{ and } \sigma_{M(i)}^2$$

[0037] are the noise energy and the masking energy associated with the critical band i . $R(i)$ is the bit rate to minimize the segmental NMR. In an $R(i)$ bits/sample PCM coder, the quantization error variance is given by

$$N(i) = \rho 2^{-2R(i)} \sigma_{x(i)}^2 \quad (2)$$

[0038] So, the minimization

$$\arg \min_{R(i)} \sum_i \left\{ \left(\frac{\rho 2^{-2R(i)} \sigma_{x(i)}^2}{\sigma_{M(i)}^2} \right) \right\} \quad (3)$$

[0039] should be constrained by the total bit rate; that is,

$$\sum_i \{R(i)B(i)\} = R. \quad (4)$$

[0040] According to the method of Lagrange multipliers, the solution must satisfy

$$\frac{\partial}{\partial R(j)} \left\{ \left(\sum_i \{R(i)B(i)\} - R \right) + \lambda \sum_i \left\{ \left(\frac{\rho 2^{-2R(i)} \sigma_{x(i)}^2}{\sigma_{M(i)}^2} \right) \right\} \right\} = 0, \text{ for all } j. \quad (5)$$

$$\text{Then } \lambda = \frac{B(j)}{(2 \log 2) \left(\frac{\rho 2^{-2R(j)} \sigma_{x(j)}^2}{\sigma_{M(j)}^2} \right)} = \frac{B(j)}{2 \log 2 \left(\frac{\sigma_{N(j)}^2}{\sigma_{M(j)}^2} \right)}, \text{ for all } j.$$

[0041] So, $R(j)$ should be allocated so that the noise-to-masking ratio is proportional to the $B(j)$. That is

$$\sigma_{N(j)}^2 = \kappa \sigma_{M(j)}^2 B(j), \text{ for all } j. \quad (6)$$

[0042] The noise level should be kept proportional to the masking threshold multiplied by a bandwidth to have the best segmental NMR.

[0043] Secondly, the noise level for the quantization bands is selected in consideration of the masking threshold and critical bandwidth in the quantization band. In other words, the

$$\sigma_{N(q)}^2$$

[0044] instead of the

$$\sigma_{N(j)}^2$$

[0045] is to be found to minimize the segmental NMR

$$\sigma_{N(q)}^2 = \kappa \sigma_{M(j)}^2 B(q) \quad (7)$$

[0046] where q is the index of the quantization band. The problem is equivalent to finding $B(q)$ to approximate best the energy defined to minimize the segmental NMR; that is

$$\hat{B}(q) = \arg \min_{B(q)} \sum_{j \in q} |\sigma_{N(q)}^2 - \sigma_{N(j)}^2| \quad (8)$$

[0047] Assume that the masking energies of the critical bands in the quantization bands are uniform, the selection after calculation is

$$\hat{B}(q) = \text{Average}(B(j))_{j \in q} \quad (9)$$

[0048] Thirdly, to avoid the bits allocated to the bands with masking level higher than the noise level, the criteria to minimize the segmental NMR is modified so that the bands with negative NMR should be rounded to 1. That is, the quantization noise for each band should have a lower bound. On the other hand, the noise higher than the masking threshold leads to a phenomenon that the associated band will be rounded to zero, referred to as the zero bands. The zero bands are quite perceptually noticeable. So, the quantization levels should also be restricted to be no larger than the signal energy.

[0049] To summarize, the bit allocation should be assigned with noise parallel to the multiplication between masking level and bandwidth under the constraints from the zero band and negative NMR.

[0050] Noise Predictor

[0051] An MPEG Layer 3 quantizer is taken as an example for the derivation of the noise predictor. From MPEG Layer 3 standard, the simplified formula for the non-uniform quantizer of layer 3 is

$$is_i = \text{int} \left(\frac{xr_i^{\frac{3}{4}}}{\Delta_q} \right), \quad (10)$$

[0052] where the quantization step size is

$$\Delta_{sfb} = 2^{\frac{3}{4}(\text{gain}_{gi} - \text{scale}_{sfb})}. \quad (11)$$

[0053] From the MPEG standard, the formula of the non-uniform quantizer can also be expressed as

$$is_i = \text{int}(xr_i 2^{\text{scale}_q - \text{gain}_{gr}} - 0.0946)^{\frac{3}{4}}, \quad (12)$$

[0054] where the scale factor is $\text{scale}_q = \frac{1}{2}(1 + \text{scalefac_scale})(\text{scalefac}_q + \text{preflag_pretab}_q)$ for each quantization band q ; scalefac_scale is 0 or 1, scalefac_q is in the range of 0~15, and the pre-amplified flag is $\text{preflag}_{gr} \cdot \text{pretab}_q$; the global gain is $\text{gain}_{gr} = \frac{1}{2}(\text{global_gain}_{gr} - 210)$ for each granule of MPEG layer 3 frame. By ignoring 0.0946, (12) can be derived as

$$\begin{aligned} is_i &= \text{int}(xr_i 2^{\text{scale}_q - \text{gain}_{gr}})^{\frac{3}{4}} \\ &= \text{int} \left(xr_i^{\frac{3}{4}} 2^{\frac{3}{4}(\text{scale}_q - \text{gain}_{gr})} \right) \\ &= \text{int} \left(\frac{xr_i^{\frac{3}{4}}}{\Delta_q} \right) \end{aligned} \quad (13)$$

[0055] where step size is

$$\Delta_q = 2^{\frac{3}{4}(\text{gain}_{gi} - \text{scale}_q)}.$$

[0056] Next, the input signal xr_1 and reconstructed signal xr_1 have the following two formulae:

$$xr_i = ((is_i + \varepsilon_i) \Delta_{sfb})^{\frac{4}{3}}, \text{ and } xr_i = (is_i \Delta_{sfb})^{\frac{4}{3}}.$$

[0057] The quantization error of the non-uniform quantizer e_1 will be equal to the difference of input signal xr_1 and reconstructed signal xr_1 :

$$e_i = xR_i - \hat{x}R_i = ((is_i + \varepsilon_i)\Delta_{gfd})^{\frac{4}{3}} - (is_i\Delta_{gfd})^{\frac{4}{3}} \quad (14)$$

$$= (1 + is_i^{-1}\varepsilon_i)^{\frac{4}{3}} is_i^{\frac{4}{3}} \Delta_{gfd}^{\frac{4}{3}} - (is_i\Delta_{gfd})^{\frac{4}{3}}$$

[0058] Let $f(\epsilon_i) = (1 + is_i^{-1}\epsilon_i)^{4/3}$. By Tyler expansion with the first order approximation of $f(\epsilon) \approx 1 + f'(\epsilon)\epsilon$, this leads to

$$e_i = f(\varepsilon_i) is_i^{\frac{4}{3}} \Delta_{gfd}^{\frac{4}{3}} - (is_i\Delta_{gfd})^{\frac{4}{3}} \approx \frac{4}{3} is_i^{\frac{1}{3}} \varepsilon_i \Delta_{gfd}^{\frac{4}{3}}.$$

[0059] Assume that the quantized signals is, and the quantized error of the uniform quantizer ϵ_1 are independent, the expectation of the quantization error of the non-uniform quantizer e , is as follows:

$$E[e_i^2] \approx \frac{16}{9} \Delta_q^{\frac{8}{3}} E[is_i^{\frac{2}{3}} \varepsilon_i^2] \approx \frac{16}{9} \Delta_q^{\frac{8}{3}} E[is_i^{\frac{2}{3}}] E[\varepsilon_i^2] \quad (15)$$

[0060] If the spectrum of the quantization bands is uniform, the noise of lines can be the average energy of quantization band; that is

$$E(e_1^2) \approx E(e_q^2) \quad (16)$$

[0061] Since

$$E[\varepsilon_i^2] = \frac{1}{12},$$

[0062] (15) becomes

$$E[e_i^2] \approx \frac{4}{27} \Delta_q^{\frac{8}{3}} E\left[\left(\frac{XR_i^{\frac{3}{4}}}{\Delta_q}\right)^{\frac{2}{3}}\right] = \frac{4}{27} \Delta_q^2 E[|XR_i|^{\frac{1}{2}}] \quad (17)$$

[0063] Substituting (7) into (16) yields

$$E[e_i^2] = \kappa \sigma_{M(q)}^2 B(q) \quad (18)$$

[0064] Finally, by defining

$$T_q = \sigma_{M(q)}^2 B(q),$$

[0065] the difference between the global gain and the scale factor is approximate to

$$gain_{gr} - scale_q \approx \frac{2}{3} \log_2 \frac{27}{4} \kappa \cdot T_q^2 / E[|XR_i|^{0.5}], \quad (19)$$

or

$$gain_{gr} - scale_q = \frac{2}{3} \left(\log_2 \frac{27}{4} + \log_2 \kappa + \log_2 T_q^2 - \log_2 E[|XR_q|^{0.5}] \right)$$

[0066] Since the scale factor $scale_q$ is in the range of 0~16 and the minimum scale for these quantization bands must be zero, thus the global gain is

$$gain_{gr} = \text{Max}_q \{ gain_{gr} - scale_q \},$$

[0067] and the scale factors for all sub-bands are obtained. It can be seen that the global gain varies with the bit rate related constant K , and the scale factor varies for each sub-band according to the masking threshold and the input signals.

[0068] Bounds on Scale Factors

[0069] As mentioned before, the bits should be allocated under non-negative NMR and the constraint of zero bands. For the non-negative NMR issues, the noise level is set to be the masking threshold; that is

$$T_q = \sigma_{M(q)}^2$$

[0070] and $K=1$. This yields to the upper bound of the U scale_q relative to the global scale.

$$gain_{gr} - U_{scale_q} = \frac{2}{3} \left(\log_2 \frac{27}{4} + \log_2 \sigma_{M(q)}^2 - \log_2 E[|XR_q|^{0.5}] \right) \quad (20)$$

That is,

$$scale_q \leq U_{scale_q} = gain_{gr} - \frac{2}{3} \left(\log_2 \frac{27}{4} + \log_2 \sigma_{M(q)}^2 - \log_2 E[|XR_q|^{0.5}] \right) \quad (21)$$

[0071] The gain_{gr} will be adjusted according to the available bits.

[0072] The lower bounds can be derived under the constraint of the zero bands. The zero bands occur when the noise is greater than the signal energy; that is

$$\Delta_q^2 = \left(2^{\frac{3}{4}(gain_{gr} - D_{scale_q})} \right)^2 < \{ E[|XR_q|^{0.5}] \}^{\frac{3}{4}} \quad (22)$$

[0073] Thus, the lower bound on the scale will be

$$scale_q \geq Dscale_q = gain_{gr} - \frac{1}{2} \log_2 E[|XR_q|^{0.5}] \quad (23)$$

[0074] FIG. 5 illustrates the average iteration number with different testing material for the present invention and the MPEG bit allocation process respectively, where Q is the quality-controlling iterations and R is the rate-controlling iterations. As shown in FIG. 5, the allocation method of the present invention has removed the iterations required for the quality-controlling iteration and have reduced the rate controlling iterations by a factor more than three.

[0075] FIG. 6 illustrates the objective score of the method of the invention compared to the bit allocation method in ISO. Here the invention adopts PEAQ (perceptual evaluation of audio quality) system which is the recommendation system by ITU-R Task Group 10/4. ISO is the original source code. ISO1 is improved by adopting the termination condition used in Lame. The experiment is based on the stereo mode and the psychoacoustic model 2. Also, since the MS switch and bit reservoir are not related to the bit allocation method, the two mechanisms have been turned off in the experiment. The objective difference grade (ODG) is the output variable from the objective measurement method. The ODG values should ideally range from 0 to -4, where 0 corresponds to an imperceptible impairment and -4 to an impairment judged as very annoying. As shown in FIG. 6, the quality from the method of the present invention is better than the suggested method in the draft.

[0076] The configuration adopted in this invention for PEAQ is the basic version. The basic version uses the FFT-based ear model. It uses the following model output variables: BandwidthRef_B, BandwidthTest_B, Total NMR_B, WinModDiff1_B, ADB_B, EHS_B, AvgModDiff1_B, AvgModDiff2_B, RmsNoiseLoud_B, MFPD_B and RelDist-Frames_B. These 11 model output variables are mapped to a single quality index using an artificial neural network with three nodes in the hidden layer.

[0077] FIG. 7 provides a list with a subset of test signals that were used during the objective and subjective test. By setting the same iteration termination conditions like iteration number, the non-increasing noise scale factor bands, fitting to scale factor table, etc [website <http://www.mp3dev.org/mp3/>], the ISO algorithm can be improved by the method mentioned in Lame (which is generally referred to as the mp3 encoder with best quality). The two nested loops adopted for the comparison is based on the iteration algorithm used in Lame.

[0078] Although the present invention has been described with reference to the preferred embodiments, it will be understood that the invention is not limited to the details described thereof. Various substitutions and modifications have been suggested in the foregoing description, and others will occur to those of ordinary skill in the art. Therefore, all such substitutions and modifications are intended to be embraced within the scope of the invention as defined in the appended claims.

What is claimed is:

1. A method of digital coding for transmitting and packing audio signals, comprising the steps of:

- (a) mapping input audio signals into a sequence of frequency samples representing a spectral composition of said audio signals;
- (b) quantizing said sequence of frequency samples into quantized values in accordance with a bit allocation process, said bit allocation process using a parameter predictor for evaluating quantization parameters by referring to a masking threshold;
- (c) encoding said quantized values using a symbol encoder to form encoded data comprising a number of bits; and

(d) packing said encoded data into a sequence of data according to a specified audio protocol.

2. The method of digital coding for transmitting and packing audio signals as claimed in claim 1, wherein said step (b) is performed either through a uniform quantizer or a non-uniform quantizer.

3. The method of digital coding for transmitting and packing audio signals as claimed in claim 1, wherein said symbol encoder comprises a VLC encoder.

4. The method of digital coding for transmitting and packing audio signals as claimed in claim 1, wherein said parameter predictor in said bit allocation process uses a deterministic formula based on a constant masking-to-noise ratio to calculate and adjust at least one corresponding global factor and/or one band scaling factor for a quantization band.

5. The method of digital coding for transmitting and packing audio signals as claimed in claim 4, wherein said bit allocation process in said step (b) further comprises the steps of adjusting said global factor according to a prescribed number of bits available for said encoded data, and yielding an upper bound and a lower bound of said band scaling factor corresponding to said global factor for a quantization band.

6. The method of digital coding for transmitting and packing audio signals as claimed in claim 5, wherein said upper bound is constrained by a non-negative noise-to-masking ratio.

7. The method of digital coding for transmitting and packing audio signals as claimed in claim 5, wherein said lower bound is constrained by zero bands.

8. The method of digital coding for transmitting and packing audio signals as claimed in claim 4, wherein said band scaling factor varies for each sub-band according to said masking threshold and said input audio signals.

9. The method of digital coding for transmitting and packing audio signals as claimed in claim 4, wherein said global factor varies with a bit rate related constant.

10. The method of digital coding for transmitting and packing audio signals as claimed in claim 1, further having an iterative rate control loop before said step (d), said iterative rate control loop comprising the steps of:

- (c1) continuing said step (d) if said number of bits comprised in said encoded data does not exceed a prescribed number of bits available for said encoded data, otherwise continuing step (c2);
- (c2) adjusting quantization parameters and a quantization step size to be used in step (b), and returning to step (b).

11. The method of digital coding for transmitting and packing audio signals as claimed in claim 10, wherein said step (b) is performed either through a uniform quantizer or non-uniform quantizer.

12. The method of digital coding for transmitting and packing audio signals as claimed in claim 10, wherein if said number of bits comprised in said encoded data exceeds a prescribed number of bits available for said encoded data, then at least one corresponding global factor and one band scaling factor are adjusted and said quantization step size is increased in said step (c2).

13. The method of digital coding for transmitting and packing audio signals as claimed in claim 10, wherein said symbol encoder comprises a VLC encoder.

14. The method of digital coding for transmitting and packing audio signals as claimed in claim 10, wherein said step (b) further comprises a step of cutting off high frequency for a low bit-rate audio coding before quantizing said sequence of frequency samples.

15. The method of digital coding for transmitting and packing audio signals as claimed in claim 14, wherein said step (c2) of said iterative rate control loop further includes adjusting a cut-off frequency for said step of cutting off high frequency.

16. The method of digital coding for transmitting and packing audio signals as claimed in claim 10, wherein said parameter predictor in said bit allocation process uses a deterministic formula based on a constant masking-to-noise ratio to calculate and adjust at least one corresponding global factor and/or one band scaling factor for a quantization band.

17. The method of digital coding for transmitting and packing audio signals as claimed in claim 16, wherein said bit allocation process in said step (b) further comprises the steps of adjusting said global factor according to a prescribed number of bits available for said encoded data, and yielding an upper bound and a lower bound of said band scaling factor corresponding to said global factor for a quantization band.

18. The method of digital coding for transmitting and packing audio signals as claimed in claim 17, wherein said upper bound is constrained by a non-negative noise-to-masking ratio.

19. The method of digital coding for transmitting and packing audio signals as claimed in claim 17, wherein said lower bound is constrained by zero bands.

20. The method of digital coding for transmitting and packing audio signals as claimed in claim 16, wherein said band scaling factor varies for each sub-band according to said masking threshold and said input audio signals.

21. The method of digital coding for transmitting and packing audio signals as claimed in claim 16, wherein said global factor varies with a bit rate related constant.

22. An architecture of digital coding for transmitting and packing audio signals, comprising:

a mapper transforming input audio signals into a sequence of frequency samples representing a spectral composition of said audio signals;

a parameter predictor evaluating quantization parameters by referring to a masking threshold;

a quantizer quantizing said sequence of frequency samples into quantized values in accordance with said quantization parameters;

a variable length encoder encoding said quantized values into encoded data comprising a number of bits; and

a packing unit packing said encoded data into a sequence of data according to a specified audio protocol.

23. The architecture of digital coding for transmitting and packing audio signals as claimed in claim 22, further comprising:

a comparator comprising said number of bits comprised in said encoded data with a prescribed number of bits available for said encoded data; and

an adjustor for adjusting said quantization parameters when said number of bits comprised in said encoded data exceeds said prescribed number of bits available for said encoded data.

24. The architecture of digital coding for transmitting and packing audio signals as claimed in claim 23, further comprising a high frequency cut-off unit connected between said mapper and said quantizer, said high frequency cut-off unit having an input for receiving a cut-off frequency from said adjustor.

* * * * *