

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
31 July 2003 (31.07.2003)

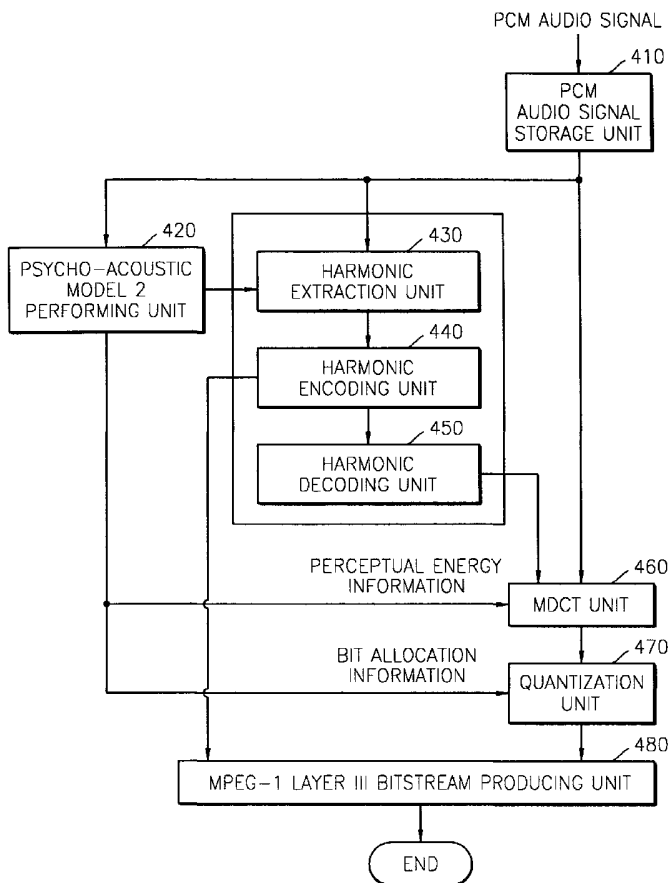
PCT

(10) International Publication Number  
WO 03/063135 A1

- (51) International Patent Classification<sup>7</sup>: **G10L 19/00**
- (21) International Application Number: PCT/KR02/02348
- (22) International Filing Date:  
12 December 2002 (12.12.2002)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
2002/36310 27 June 2002 (27.06.2002) KR
- (71) Applicant: **SAMSUNG ELECTRONICS CO., LTD.**  
[KR/KR]; 416, Maetan-dong, Paldal-gu, Suwon-shi,  
Kyunagki-do 442-373 (KR).
- (72) Inventor: **HA, Ho-Jin**; (21/2) 994-39 Doksan 3-dong,  
Geumcheon-gu, Seoul 153-013 (KR).
- (74) Agent: **LEE, Young-Pil**; The Cheonghwa Building,  
1571-18, Seocho-dong, Seocho-gu, Seoul 137-874 (KR).
- (81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK,  
LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX,  
MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG,  
SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN,  
YU, ZA, ZM, ZW.
- (84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: AUDIO CODING METHOD AND APPARATUS USING HARMONIC EXTRACTION



(57) Abstract: A method and apparatus for effectively encoding an audio signal into a Moving Picture Experts Group (MPEG)-1 layer III audio signal of a low-speed bitrate. In the audio encoding method, harmonic components are extracted using fast Fourier transformation (FFT) result information that is obtained by applying psycho-acoustic model 2 to received pulse code modulation (PCM) audio data. Then, the extracted harmonic components are removed from the received PCM audio data. Thereafter, the PCM audio data from which the extracted harmonic components are removed is subjected to a modified discrete cosine transform (MDCT) and quantization. Accordingly, efficient encoding can be achieved even using a small number of allocated bits.



WO 03/063135 A1



ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

— upon request of the applicant, before the expiration of the time limit referred to in Article 21(2)(a)

**Published:**

- with international search report
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

## AUDIO CODING METHOD AND APPARATUS USING HARMONIC EXTRACTION

### Technical Field

5           The present invention relates to a method of compressing an audio signal, and more particularly, to a method of and apparatus for efficiently compressing an audio signal into an MPEG-1 layer-3 audio signal with a low-speed bit rate.

### 10   Background Art

          Moving Picture Experts Group-1 (MPEG-1) establishes standards regarding digital video compression and digital audio compression and is supported by the International Standardization Organization (ISO). The MPEG-1 audio standard is used to compress 16-bit audio that is  
15   sampled at a 44.1Khz sampling rate and stored on a 60-minute or 72-minute CD and is classified into 3 layers according to a compression method and the complexity of a codec.

          Layer III is the most complex, uses much more filters than layer II, and adopts Huffman coding. Upon encoding at 112 Kbps, an  
20   excellent-quality sound can be heard. Upon encoding at 128 Kbps, a sound nearly the same as the original sound is obtained. Upon encoding at 160 Kbps or 192 Kbps, an excellent sound that a human ear cannot distinguish from the original sound can be heard. In general, MPEG-1 layer-3 audio is referred to as MP3 audio.

25           MP3 audio is produced through a discrete cosine transform (DCT), bit allocation based on psycho-acoustic model 2, quantization, and the like. To be more specific, while the number of bits used to compress audio data is kept to be minimal, modified DCT (MDCT) is performed using the result of psycho-acoustic model 2.

In audio compression techniques, the ear of a human is the most important. The human ear cannot hear if the intensity of a sound is at or below a predetermined level. If someone talks loudly in the office, it can be easily recognized who is talking. However, if an airplane  
5 passes at that moment, the talking cannot be heard. Even after the airplane has passed, the talking still cannot be heard because of a lingering sound. Accordingly, in psycho-acoustic model 2, data having a volume equal to or greater than a masking threshold is sampled among data having a volume equal to or greater than the minimum  
10 audible limit corresponding to when it is quiet. The sampling is performed on each sub-band.

However, when a sound signal is compressed at a low-speed bit rate of no more than 64Kbps, psycho-acoustic model 2 is not suitable because the number of bits used to quantize a signal such as a pre-echo  
15 signal is limited. Hence, in order to overcome this problem caused by low-speed MP3 audio, the present invention provides a method of effectively processing an audio signal at a low speed by removing a harmonic component from an original signal using a fast Fourier transform (FFT) adopted in psycho-acoustic model 2 and compressing  
20 only a transient component using MDCT.

In an FFT process adopted in a conventional psycho-acoustic model, only signal analysis is performed, and the result of the FFT is not used. Since the result of the FFT is not used for signal compression, it can be considered to be a waste of resources.

25 Korean Patent Publication No. 1995-022322 discloses a bit allocation method employing a psycho-acoustic model. However, the disclosed method is different from a method of the present invention for increasing compression efficiency by removing a harmonic component from an original signal using the result of an FFT adopted in a  
30 psycho-acoustic model.

Korean Patent Publication No. 1998-072457 discloses a signal processing method and apparatus in the psycho-acoustic model 2, by which the amount of computation is significantly reduced by reducing computation overload while compressing an audio signal. That is, the disclosed signal processing method includes a step of obtaining an individual masking boundary value using an FFT result, a step of selecting a global masking boundary value, and a step of shifting to the next frequency position. This method is the same as the present invention in that an FFT result value is used but different in that it uses a different quantization method.

U.S. Patent No. 5,930,373 discloses a method for enhancing the quality of a sound signal using the residue harmonics of a low frequency signal. However, the disclosed method and the quantization method according to the present invention are different in that they use different techniques of using residue harmonics.

#### Disclosure of the Invention

To solve the above and other problems, it is an aspect of the present invention to provide a method of effectively processing an audio signal at a low speed by removing a harmonic component from an original audio signal using the result of a fast Fourier transform (FFT) used in psycho-acoustic model 2 and compressing only a residue transient using a modified discrete cosine transform (MDCT).

The above and other aspects of the present invention are achieved by an audio coding method using harmonic components. In this method, first, pulse code modulation (PCM) audio data are received, and harmonic components are extracted from the received PCM audio data by applying psycho-acoustic model 2. Next, a modified discrete cosine transform (MDCT) is performed on the received PCM audio data from which the extracted harmonic components are removed.

Thereafter, the MDCTed audio data is quantized, and an audio packet is produced from quantized audio data and the extracted harmonic components.

The above and other aspects of the present invention are also  
5 achieved by an audio coding method using harmonic components, in which PCM audio data is first received and stored. Then, psycho-acoustic model 2 based on the audible limit characteristics of a human is applied to the stored data to obtain fast Fourier transformation (FFT) result, perceptual energy information regarding received data, and  
10 bit allocation information used for quantization. Thereafter, harmonic components are extracted from the received PCM audio data using the FFT result information. Next, the extracted harmonic components are encoded, and the encoded harmonic components are decoded. Then, a MDCT is performed on a number of samples of the received PCM  
15 audio data from which the extracted harmonic components are removed, which depends on the value of the perceptual energy information. Thereafter, the MDCTed audio data is quantized by allocating bits according to the bit allocation information. Finally, an audio packet is produced from the quantized, MDCTed audio data and the encoded  
20 harmonic components.

The above and other aspects of the present invention are still achieved by an audio coding apparatus using harmonic components. In the apparatus, a PCM audio data storage unit receives and stores PCM audio data. A psycho-acoustic model 2 performing unit receives the  
25 PCM audio data from the PCM audio data storage unit and performs psycho-acoustic model 2 to obtain FFT result information, perceptual energy information regarding received data, and bit allocation information used for quantization. A harmonic extraction unit extracts harmonic components from the received PCM audio data using the FFT  
30 result information. A harmonic encoding unit encodes the extracted

harmonic components outputting encoded harmonic components. A harmonic decoding unit decodes the encoded harmonic components. An MDCT unit performs a MDCT on the stored PCM audio data from which the decoded harmonic components are removed, according to the perceptual energy information. A quantization unit quantizes the MDCTed audio data according to the bit allocation information. An MPEG layer III bitstream production unit transforms the quantized, MDCTed audio data and the encoded harmonic components output from the harmonic encoding unit into an MPEG audio layer III packet.

To achieve the above and other aspects, the present invention provides a computer readable recording medium which stores a computer program for executing the above methods.

#### Brief Description of the Drawings

FIG. 1 shows the format of an MPEG-1 layer III audio stream;

FIG. 2 is a block diagram of an apparatus for producing an MPEG-1 layer III audio stream;

FIG. 3 is a flowchart illustrating a computation process in a psycho-acoustic model;

FIG. 4 is a block diagram of an apparatus according to the present invention for producing a low-speed MPEG-1 layer III audio stream;

FIG. 5 is a flowchart illustrating harmonic extraction, harmonic encoding, and harmonic decoding based on psycho-acoustic model 2;

FIGS. 6A, 6B, 6C, and 6D illustrate harmonic component samples extracted in stages in order to extract harmonic components using an FFT result in psycho-acoustic model 2;

FIG. 7 is a table showing limited frequency ranges varying according to K values; and

FIG. 8 is a flowchart illustrating a process according to the present invention for producing an audio stream by removing a harmonic component.

5 Best mode for carrying out the Invention

Referring to FIG. 1, a moving picture experts group (MPEG)-1 layer III audio stream is composed of audio access units (AAUs) 100. The AAU 100 is a minimal unit that can be independently accessed, and compresses and stores data with a fixed number of samples. The AAU  
10 100 includes a header 110, a cyclic redundancy check (CRC) 120, audio data 130, and auxiliary data 140.

The header 110 stores a syncword, ID information, layer information, information regarding whether a protection bit exists, bitrate index information, sampling frequency information, information regarding  
15 whether a padding bit exists, a private bit, mode information, mode extension information, copyright information, information regarding whether an audio stream is an original one or a copy, and information on emphasis characteristics.

The CRC 120 is optional. The presence or absence of the CRC  
20 120 is defined in the header 110, and the length of the CRC 120 is 16 bits.

The audio data 130 is a portion into which compressed audio data is inserted.

The auxiliary data 140 is data which is filled into a space  
25 remaining when the end of the audio data 130 does not reach the end of an AAU. Arbitrary data other than MPEG audio can be inserted into the auxiliary data 140.

FIG. 2 is a block diagram of an apparatus for producing an MPEG-1 layer III audio stream. A pulse code modulation (PCM) audio  
30 signal input unit 210 has a buffer in which PCM audio data is stored.

Here, the PCM audio signal input unit 210 receives, as the PCM audio data, granules, each composed of 576 samples.

A psycho-acoustic model 2 performing unit 220 receives the PCM audio data from the buffer of the PCM audio signal input unit 210 and performs psycho-acoustic model 2. A discrete cosine transforming (DCT) unit 230 receives the PCM audio data in units of granules and performs a DCT operation at the same time when psycho-acoustic model 2 is performed.

A modified DCT (MDCT) unit 240 performs an MDCT using the result of the application of psycho-acoustic model 2 and the result of the DCT performed by the DCT unit 230. If perceptual energy is greater than a predetermined threshold, the MDCT is performed using a short window. If the perceptual energy is smaller than the predetermined threshold, the MDCT is performed using a long window.

In perceptual coding, which is an audio signal compression technique, a reproduced signal is different to an original signal. That is, detailed information that people cannot perceive using the characteristics of the human ear can be omitted. Perceptual energy denotes energy that a human can perceive.

A quantization unit 250 performs quantization using bit allocation information generated as a result of the application of psycho-acoustic model 2 and using the result of the MDCT operation. An MPEG-1 layer III bitstream producing unit 260 transforms the quantized data into data to be inserted into an audio data area of an MPEG-1 bitstream, using Huffman coding.

FIG. 3 is a flowchart illustrating a computation process in a psycho-acoustic model. First, PCM audio data is received in granules, each composed of 576 samples, in step 310. Next, long windows, each composed of 1024 samples, or short windows, each composed of 256

samples, are formed using the received PCM audio data, in step 320. That is, one packet is constituted of multiple samples.

Thereafter, in step 330, a fast Fourier transform (FFT) is performed one window at a time on the windows formed in step 320.

5 Then, psycho-acoustic model 2 is applied, in step 340.

In step 350, a perceptual energy value is obtained through the application of psycho-acoustic model 2 and applied to a MDCT unit and the MDCT unit selects a window to be applied. A signal to masking ratio (SMR) value for each threshold bandwidth is calculated and applied  
10 to a quantization unit to determine the number of bits to be allocated.

Finally, MDCT and quantization are performed using the perceptual energy value and the SMR value, in step 360.

FIG. 4 is a block diagram of an apparatus for producing a low-speed MPEG-1 layer III audio stream, according to the present  
15 invention. A PCM audio signal storage unit 410 has a buffer in which it stores PCM audio data. A psycho-acoustic model 2 performing unit 420 performs an FFT on 1024 samples or 256 samples at a time and outputs perceptual energy information and bit allocation information.

As described above with reference to FIG. 3, when  
20 psycho-acoustic model 2 is applied, the perceptual energy information and the bit allocation information that depends on an SMR are output. Since the psycho-acoustic model 2 performing unit 420 performs an FFT, a harmonic extraction unit 430 extracts a harmonic component from the result of the FFT. This will be described later with reference to FIG. 6.

25 A harmonic encoding unit 440 encodes the extracted harmonic component and transmits the encoded harmonic component to an MPEG-1 layer III bitstream producing unit 480. The encoded harmonic component forms MPEG-1 audio, together with quantized audio data. The encoding process of a harmonic component will be described later  
30 in detail.

A harmonic decoding unit 450 decodes the encoded harmonic component to obtain PCM data in the time domain. A MDCT unit 460 subtracts the decoded harmonic component from the original input PCM signal and performs a MDCT on the result of the subtraction. To be  
5 more specific, if the perceptual energy information value received from the psycho-acoustic model 2 unit 420 is greater than a predetermined threshold, a MDCT is performed on 18 samples at a time. If the perceptual energy information value received from the psycho-acoustic model 2 performing unit 420 is equal to or smaller than the  
10 predetermined threshold, a MDCT is performed on 36 samples at a time.

The harmonic component extraction is performed on data arranged in a frequency domain using a tonal/non-tonal decision condition and auditory limit characteristics that are defined in psycho-acoustic model 2. This will be described later in detail.

15 A quantization unit 470 performs quantization using the bit allocation information obtained by the psycho-acoustic model 2 performing unit 420. The MPEG-1 layer III bitstream producing unit 480 packetizes the harmonic component data made by the harmonic encoding unit 440 and quantized audio data obtained by the  
20 quantization unit 470 to obtain compressed audio data.

FIG. 5 is a flowchart illustrating a harmonic extraction step 510, a harmonic encoding step 520, and a harmonic decoding step 530 based on psycho-acoustic model 2. The steps performed in psycho-acoustic model 2 in FIG. 5 are the same as the steps performed in  
25 psycho-acoustic model 2 in FIG. 3. The result of the FFT performed based on the psycho-acoustic model 2 performing unit is used in step 510 of extracting a harmonic component. The extracted harmonic component is encoded to an MPEG-1 bitstream in step 520. The harmonic extraction step 510 will now be described in greater detail with  
30 reference to FIGS. 6A through 6D.

FIGS. 6A, 6B, 6C, and 6D illustrate samples extracted in stages when harmonic components are extracted using the result of the FFT performed in psycho-acoustic model 2. If PCM audio data as shown in FIG. 6A are input, an FFT is first performed on the received data in order to determine sound pressure for each datum. One of the plurality of received PCM audio data whose sound pressure has been obtained is selected. If the values of the PCM audio data on the left and right sides of the selected data are smaller than the selected PCM audio data value, only the selected PCM audio data is extracted. This process is applied to all of the received PCM audio data.

Sound pressure is the energy value of a sample in a frequency domain. In the present invention, only samples having sound pressures that are greater than a predetermined level are determined to be harmonic components. Accordingly, the samples shown in FIG. 6B are extracted. Thereafter, only samples having sound pressures that are greater than a predetermined level are extracted. For example, if the predetermined level is set to be 7.0 dB, samples having sound pressures smaller than 7.0 dB are not selected, and only the samples shown in FIG. 6C remain. The remaining samples are not all considered to be harmonic components, and some samples are extracted from the remaining samples according to the table of FIG. 7. Hence, finally, the samples shown in FIG. 6D remain.

FIG. 7 is a table showing a limited frequency range that varies according to a K value. Given that K is a value representing the location of a sample in a frequency domain, if the K value is smaller than 3 or greater than 500, the values of samples present within the limited frequency range of 0 are 0 and accordingly not selected. Likewise, as shown in FIG. 7, if the K value is equal to or greater than 3 and smaller than 63, a corresponding range value is set to be 2. If the K value is equal to or greater than 63 and smaller than 127, a corresponding range

value is set to be 3. If the K value is equal to or greater than 127 and smaller than 255, a corresponding range value is set to be 6. If the K value is equal to or greater than 255 and smaller than 500, a corresponding range value is set to be 12.

5           Setting 500 as the limit was made in consideration of the limit of the audible frequency of a human and was based on an assumption that there is no difference in the quality of reproduced sound between when sample values corresponding to a frequency equal to or greater than 500 are considered and when they are not considered.

10           Consequently, only the sample values of FIG. 6D are extracted and determined to be harmonic components.

Harmonic encoding 520 includes amplitude encoding, frequency encoding, and phase encoding. These three encoding methods use Equations 1 and 2:

$$\text{Enc\_peak\_AmpMax} = \text{integer}\left(\frac{(2^8 - 1)\log_{10}(\text{AmpMax} + 10)}{\log_{10} 2^{13}}\right) \quad \dots(1)$$

$$\text{Enc\_Amp} = \text{integer}\left(\frac{(2^5 - 1)\log_{10}(\text{Amp} + 10)}{\log_{10}(\text{AmpMax} + 10)}\right) \quad \dots(2)$$

15           wherein AmpMax denotes a peak amplitude, Enc\_peak-AmpMax denotes a result value obtained by encoding the value AmpMax, and Amp denotes amplitudes other than the peak amplitude.

          In the amplitude encoding, when a peak amplitude is set as the value AmpMax, the peak amplitude is first encoded in a 8-bit log scale to  
20           obtain Enc\_peak\_AmpMax as shown in Equation 1, and the other amplitudes Amp are encoded in a 5-bit log scale to obtain Enc-Amp as shown in Equation 2.

          In the frequency encoding, only samples corresponding to values K ranging from 58 (2498Hz) to 372 (16KHz) are encoded in  
25           consideration of the human's auditory characteristics. Since 314 is

obtained by subtracting 58 from 372, the samples are encoded using 9 bits.

The phase encoding is achieved using 3 bits.

After such harmonic extraction and harmonic encoding, encoded  
5 harmonic components are decoded and then undergo MDCT.

FIG. 8 is a flowchart illustrating a process for producing an audio stream by removing harmonic components, according to the present invention. First, in step 810, PCM audio data is received and stored. Then, in step 820, psycho-acoustic model 2 using the audible limit  
10 characteristics of a human being is applied to the stored data in order to obtain FFT result information, perceptual energy information regarding the received data, and bit allocation information used for quantization. Thereafter, in step 830, harmonic components are extracted from the received PCM audio data using the FFT result information.

15 The harmonic components are extracted in the following process.

First, sound pressure for each of the plurality of received PCM audio data is obtained using the FFT result information. Next, one of the plurality of received PCM audio data whose sound pressures are obtained is selected. If the values of the PCM audio data on the left  
20 and right sides of the selected data are smaller than the value of the selected PCM audio data, only the selected PCM audio data is extracted.

This process is applied to all of the received PCM audio data. Thereafter, only PCM audio data that each have sound pressure greater than a predetermined value of 7.0 dB are extracted from the PCM audio  
25 data extracted in the previous step. Finally, harmonic components are extracted by not selecting PCM audio data in a predetermined frequency range among the audio data extracted in the previous step.

After the harmonic extraction in step 830, the extracted harmonic components are encoded and output in step 840. Then, encoded  
30 harmonic components are decoded in step 850.

Next, in step 860, the received PCM audio data from which the decoded harmonic components are removed is subject to MDCT according to the perceptual energy information. To be more specific, if a perceptual energy value is greater than a predetermined threshold, MDCT is performed using a short window, for example, on 18 samples at a time. If the perceptual energy value is smaller than the predetermined threshold, MDCT is performed using a long window, for example, on 36 samples at a time.

Thereafter, in step 870, the MDCT result values are quantized by allocating bits according to the bit allocation information.

Finally, in step 880, the quantized audio data and the encoded harmonic components are subject to Huffman coding to obtain an audio packet.

The embodiments of the present invention can be written as computer programs and can be implemented in general-use digital computers that execute the programs using a computer readable recording medium. Examples of computer readable recording media include magnetic storage media (e.g., ROM, floppy disks, hard disks, etc.), optical recording media (e.g., CD-ROMs, or DVDs), and a storage medium such as a carrier wave (e.g., transmission through the Internet).

While the present invention has been particularly shown and described with reference to preferred embodiments thereof, it will be understood by those of ordinary skill in the art that various changes in form and details may be made therein without departing from the spirit and scope of the present invention as defined by the following claims. Hence, disclosed embodiments must be considered not restrictive but explanatory. The scope of the present invention is not presented in the above description but in the following claims, and all difference in the equivalent scope to the scope of the claims must be interpreted as being included in the present invention.

### Industrial Applicability

As described above, in the present invention, the number of quantization bits generated upon production of a low-speed MPEG-1 layer III audio stream is minimized. Using FFT results used in psycho-acoustic model 2, harmonic components are simply removed from an input audio signal, and only a transient portion is compressed using MDCT. Therefore, the input audio signal can be effectively compressed at a low-speed bitrate.

What is claimed is:

1. An audio coding method using harmonic components, the method comprising:
  - (a) receiving pulse code modulation (PCM) audio data and  
5 extracting harmonic components from the received PCM audio data by applying psycho-acoustic model 2;
  - (b) performing a modified discrete cosine transform (MDCT) on the received PCM audio data from which the extracted harmonic components are removed; and  
10 (c) quantizing the MDCTed audio data and producing an audio packet from quantized audio data and the extracted harmonic components.
  
2. An audio coding method using harmonic components, the  
15 method comprising:
  - (a) receiving and storing PCM audio data and applying psycho-acoustic model 2 based on the audible limit characteristics of a human to the stored data to obtain fast Fourier transformation (FFT)  
20 result, perceptual energy information regarding received data, and bit allocation information used for quantization;
  - (b) extracting harmonic components from the received PCM audio data using the FFT result information;
  - (c) encoding the extracted harmonic components, outputting encoded harmonic components, and decoding the encoding harmonic  
25 components;
  - (d) performing a MDCT on a number of samples of the received PCM audio data from which the extracted harmonic components are removed, which depends on the value of the perceptual energy information;

(e) quantizing the MDCTed audio data by allocating bits according to the bit allocation information; and

(f) producing an audio packet from the quantized, MDCTed audio data and the encoded harmonic components.

5

3. The audio coding method of claim 2, wherein step (b) comprises:

(b1) obtaining sound pressures for the plurality of received PCM audio data using the FFT result information;

10 (b2) selecting a data value from the plurality of PCM audio data whose sound pressure is obtained and firstly extracting only the selected PCM audio datum if the value of PCM audio data on the right and left sides of the selected PCM audio data value are smaller than the selected PCM audio data value;

15 (b3) applying step (b2) to all of the received PCM audio data;

(b4) secondly extracting only the PCM audio data whose sound pressures are greater than a predetermined sound pressure, from the firstly-extracted PCM audio data; and

20 (b5) not selecting PCM audio data that exist within a predetermined frequency range depending on a frequency location, among the PCM audio data secondly extracted in step (b4).

4. The audio coding method of claim 3, wherein the predetermined sound pressure in step b5 is 7.0 dB.

25

5. The audio coding method of claim 2, wherein in step (d), if the value of the perceptual energy information is greater than a predetermined threshold, MDCT is performed on 18 samples at a time, or if the value of the perceptual energy information is smaller than the  
30 predetermined threshold, MDCT is performed on 36 samples at a time.

6. An audio coding apparatus using harmonic components, the apparatus comprising:
- 5 a PCM audio data storage unit receiving and storing PCM audio data;
  - a psycho-acoustic model 2 performing unit receiving the PCM audio data from the PCM audio data storage unit and performing psycho-acoustic model 2 to obtain FFT result information, perceptual energy information regarding received data, and bit allocation
  - 10 information used for quantization;
  - a harmonic extraction unit extracting harmonic components from the received PCM audio data using the FFT result information;
  - a harmonic encoding unit encoding the extracted harmonic components outputting encoded harmonic components;
  - 15 a harmonic decoding unit decoding the encoded harmonic components;
  - an MDCT unit performing MDCT on the stored PCM audio data from which the decoded harmonic components are removed, according to the perceptual energy information;
  - 20 a quantization unit quantizing the MDCTed audio data according to the bit allocation information; and
  - an MPEG layer III bitstream production unit transforming the quantized, MDCTed audio data and the encoded harmonic components output from the harmonic encoding unit into an MPEG audio layer III
  - 25 packet.

7. The audio coding apparatus of claim 6, wherein the harmonic extraction unit performs harmonic extraction through the following steps of:

obtaining sound pressures for the plurality of received PCM audio data using the FFT result information, selecting a datum from the plurality of PCM audio data whose sound pressures are obtained, and firstly extracting only the selected PCM audio datum if the value of PCM audio data on the right and left sides of the selected PCM audio datum are smaller than the value of the selected PCM audio datum;

5 applying the firstly-extraction to all of the received PCM audio data and secondly extracting only the PCM audio data whose sound pressures are greater than a predetermined sound pressure, from the firstly-extracted PCM audio data; and

10 deserting PCM audio data that exist within a predetermined frequency range depending on a frequency location, from the secondly-extracted PCM audio data.

15 8. The audio coding apparatus of claim 6, wherein the MDCT unit performs MDCT on 18 samples at a time if the value of the perceptual energy information is greater than a predetermined threshold, or performs MDCT on 36 samples at a time if the value of the perceptual energy information is smaller than the predetermined threshold.

20

9. A computer readable recording medium which stores a computer program for executing claim 1.

25 10. A computer readable recording medium which stores a computer program for executing claim 2.

1/7

FIG. 1

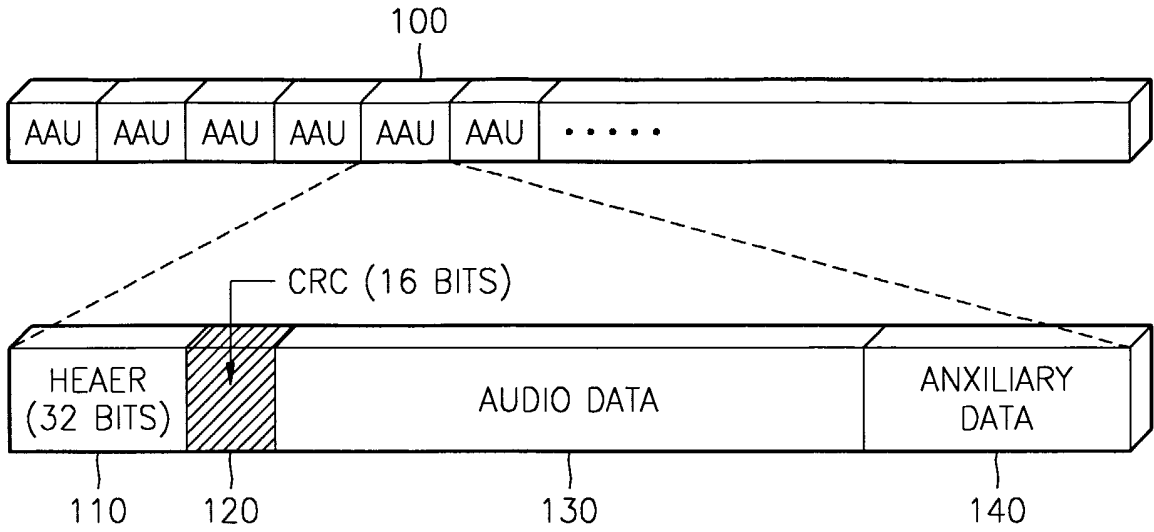
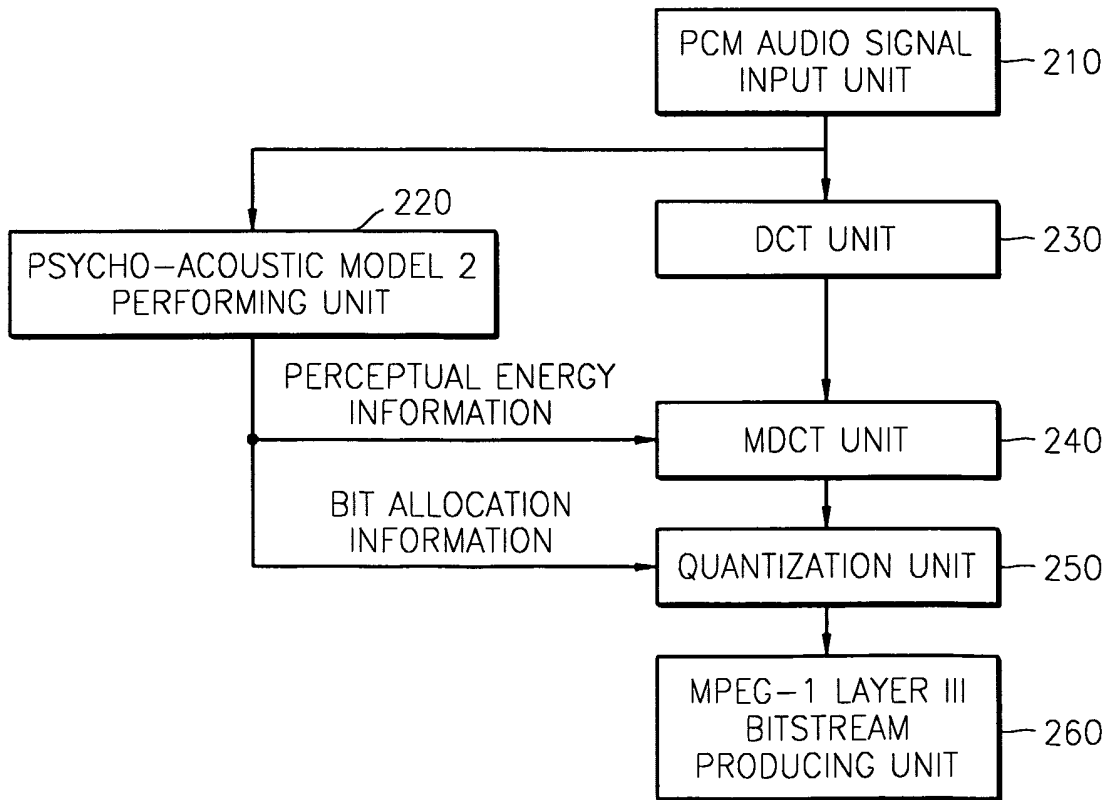
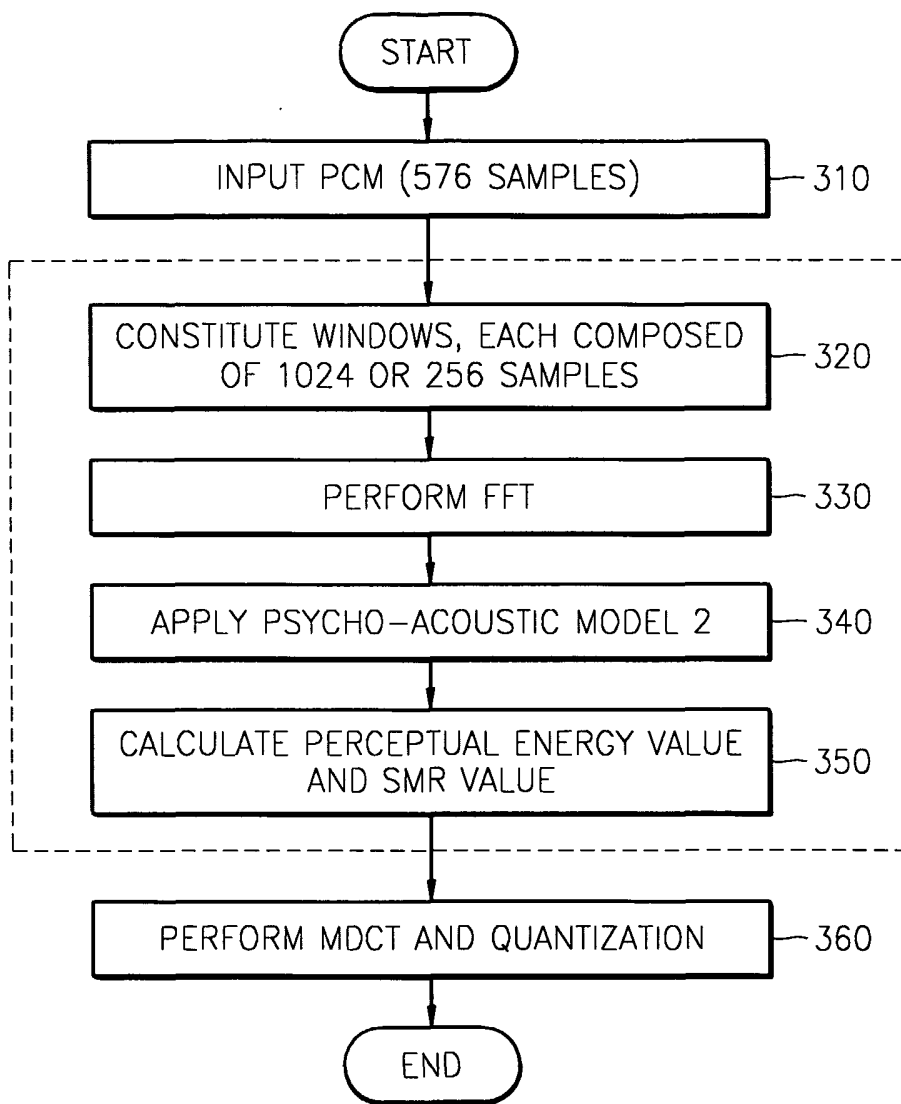


FIG. 2



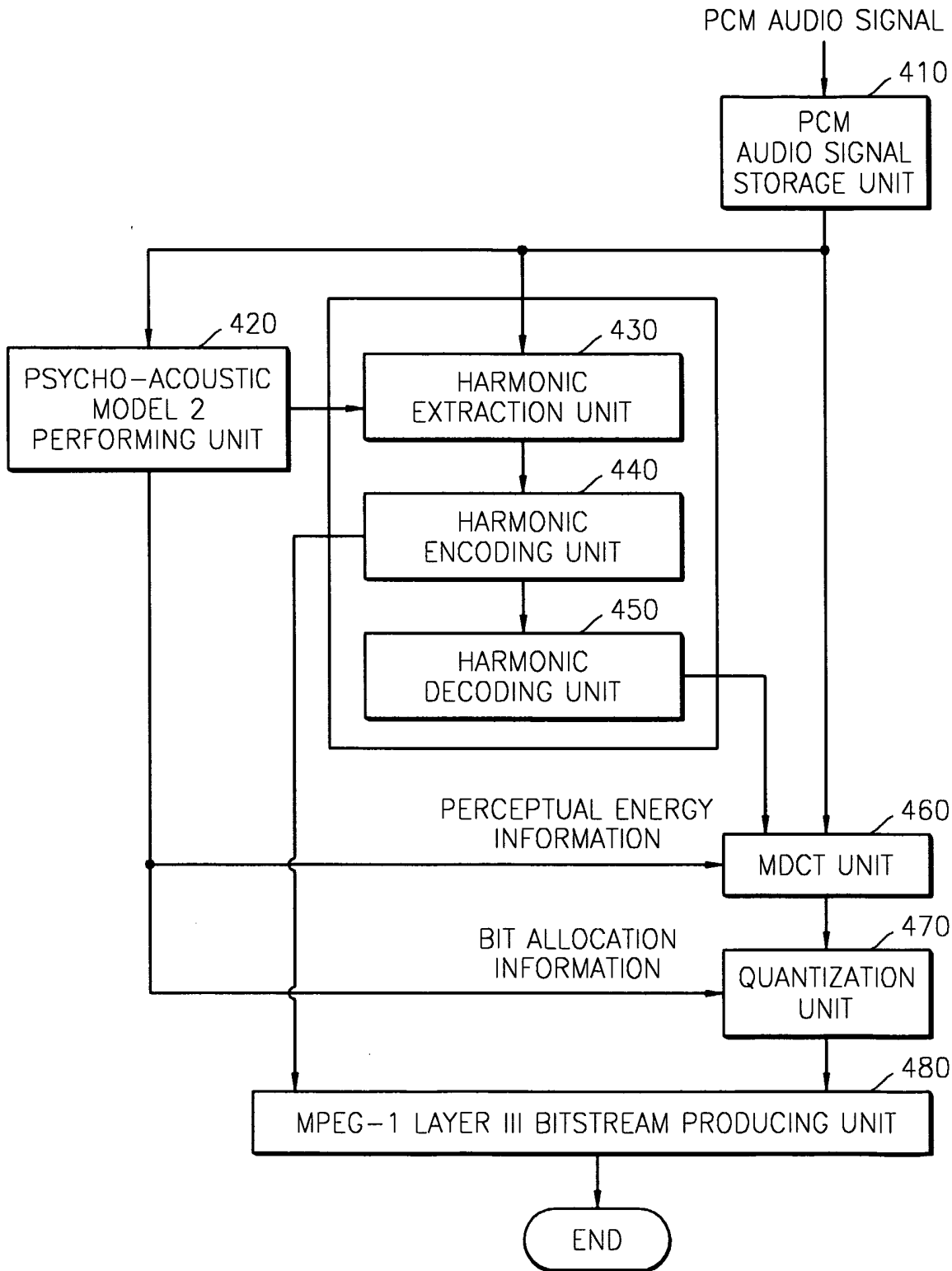
2/7

FIG. 3



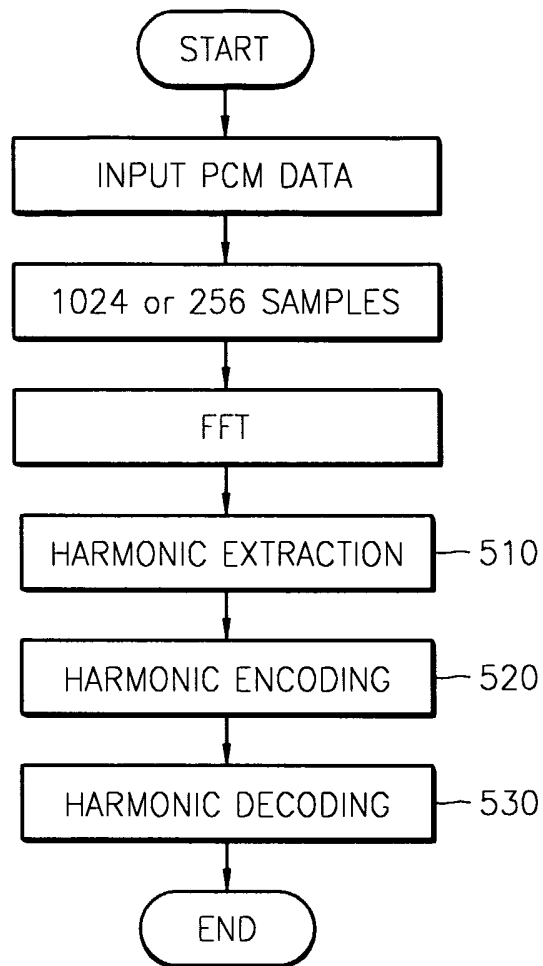
3/7

FIG. 4



4/7

FIG. 5



5/7

FIG. 6A

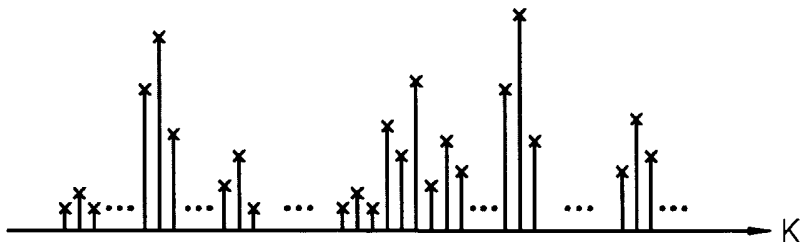


FIG. 6B

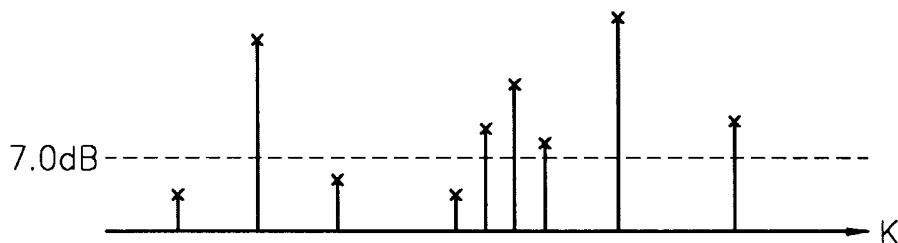


FIG. 6C

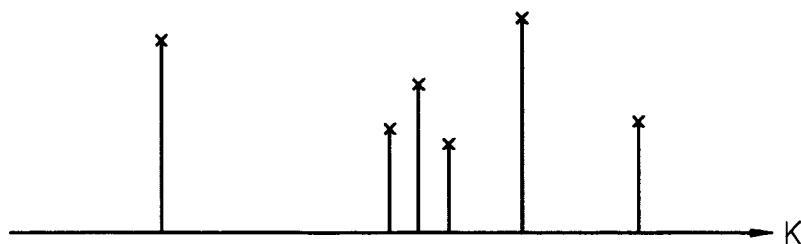
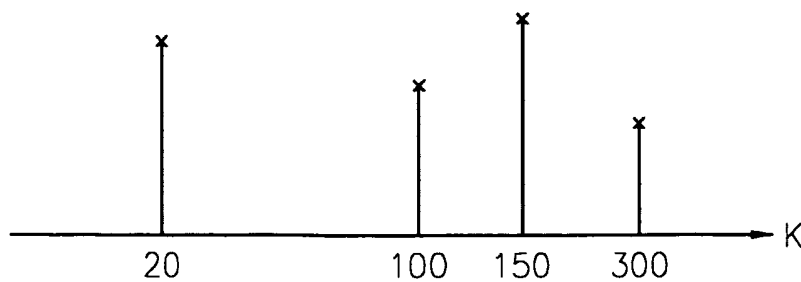


FIG. 6D



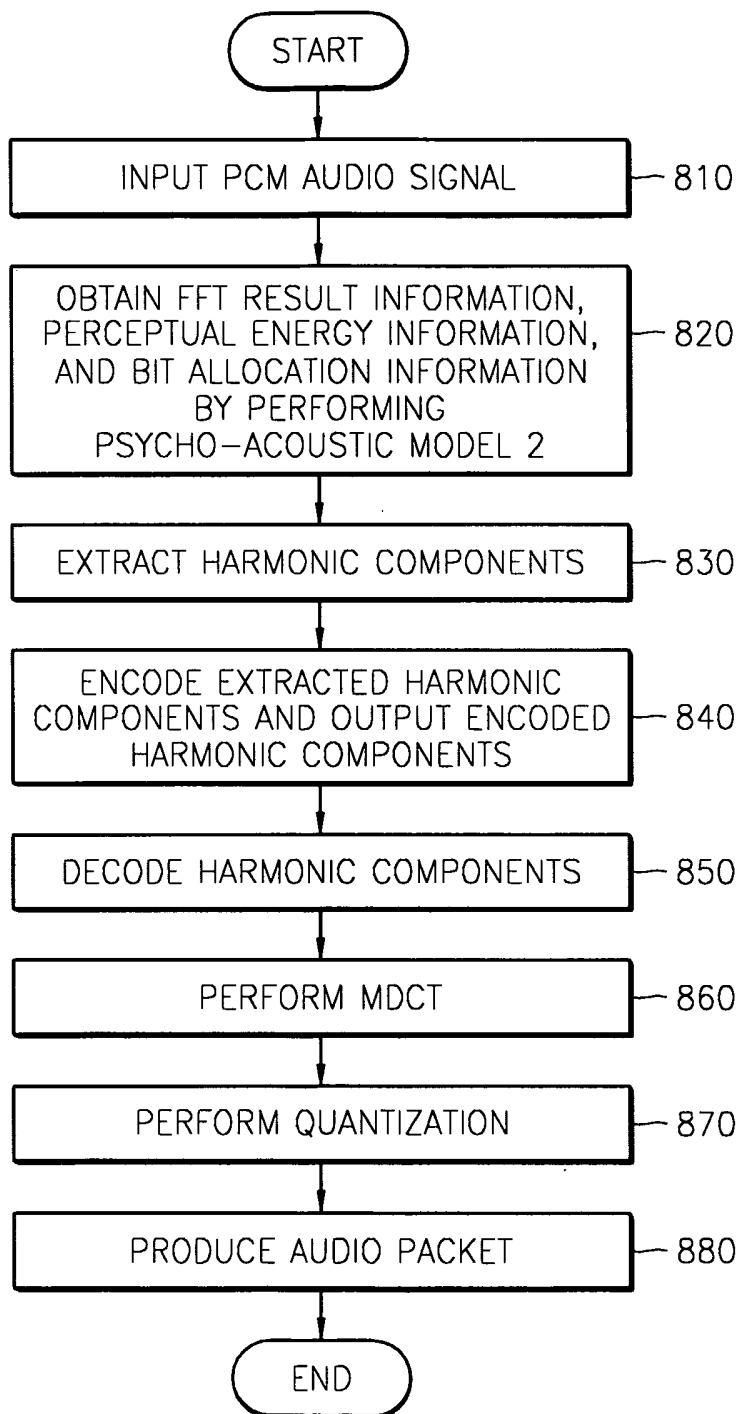
6/7

**FIG. 7**

LOCATION OF SOUND PRESSURE IN FREQUENCY DOMAIN	LIMITED FREQUENCY RANGE
$K < 3$ or $K > 500$	RANGE = 0
$3 \leq K < 63$	RANGE = 2
$63 \leq K < 127$	RANGE = 3
$127 \leq K < 255$	RANGE = 6
$255 \leq K < 500$	RANGE = 12

7/7

FIG. 8



INTERNATIONAL SEARCH REPORT

International application No.  
PCT/KR02/02348

**A. CLASSIFICATION OF SUBJECT MATTER**  
**IPC7 G10L 19/00**  
According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**  
Minimum documentation searched (classification system followed by classification symbols)  
IPC7 G10L 19/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched  
Korea Patents and applications for inventions since 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
NPS : "audio, coding, mdct, harmonic, mpeg"

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, A	KR 2002-0077959 A(LG Electronics Co.Ltd.) 18 October 2002 "the whole document"	1-10
A	KR 1998-072457 A(Dongyang Computer Co. Ltd.) 05 November 1998 "claims 1-4"	1-10
A	US 5,481,614 A(AT&T Corp.) 02 January 1996 "the whole document"	1-10
A	US 5,819,212 A(Sony Corp.) 06 October 1998 "the whole document"	1-10



Further documents are listed in the continuation of Box C.

See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 13 FEBRUARY 2003 (13.02.2003)	Date of mailing of the international search report 17 FEBRUARY 2003 (17.02.2003)
--	---

Name and mailing address of the ISA/KR  Korean Intellectual Property Office 920 Dunsan-dong, Seo-gu, Daejeon 302-701, Republic of Korea Facsimile No. 82-42-472-7140	Authorized officer NAM, In Ho Telephone No. 82-42-481-5761 
---	--

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/KR02/02348

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5,819,212 A	06. 10. 1998	EP1262956A3 JP9127986A CN1154013A	04.12.2002 16.05.1997 02.05.1997