



(51) International Patent Classification:

G10L 19/02 (2013.01) G10L 21/02 (2013.01)
G10L 19/028 (2013.01)

PCT/US2015/065936

16 December 2015 (16.12.2015)

US

(21) International Application Number:

PCT/US2016/021976

(71) Applicant: PSYX RESEARCH, INC. [US/US]; 17799
Cartwright Road, Unit 101, Irvin, California 92614 (US).

(22) International Filing Date:

11 March 2016 (11.03.2016)

(72) Inventor: REAMS, Robert W.; 688 110th Ave. NE
#1105-S, Bellevue, Washington 98004 (US).

(25) Filing Language:

English

(74) Agents: ROURK, Christopher J. et al.; Jackson Walker
L.L.P., 2323 Ross Avenue, Suite 600, Dallas, Texas 75201
(US).

(26) Publication Language:

English

(30) Priority Data:

62/133,167	13 March 2015 (13.03.2015)	US
62/156,065	1 May 2015 (01.05.2015)	US
62/156,061	1 May 2015 (01.05.2015)	US
14/863,368	23 September 2015 (23.09.2015)	US
14/863,365	23 September 2015 (23.09.2015)	US
14/863,357	23 September 2015 (23.09.2015)	US
14/863,350	23 September 2015 (23.09.2015)	US
14/863,374	23 September 2015 (23.09.2015)	US
14/863,376	23 September 2015 (23.09.2015)	US

(81) Designated States (unless otherwise indicated, for every
kind of national protection available): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY,
BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,
MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM,
PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC,
SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN,
TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

[Continued on next page]

(54) Title: SYSTEM AND METHOD FOR DYNAMIC RECOVERY OF AUDIO DATA AND COMPRESSED AUDIO ENHANCEMENT

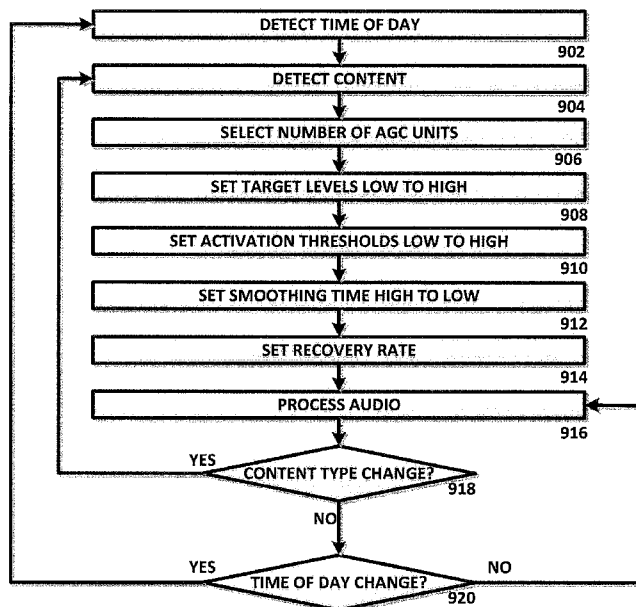


FIG. 9

900



(57) Abstract: A system for processing digitally-encoded audio data comprising a compressed audio source device providing a sequence of frames of compressed digital audio data. A compressed audio enhancement system configured to receive the sequence of frames of compressed digital audio data and to generate enhanced audio data by adding masked digital audio data to the sequence of frames of compressed digital audio data, where the masked digital audio data has an energy level sufficient to keep a kinocilia of a listener active. One or more speakers configured to receive the enhanced audio data and to generate sound waves using the enhanced audio data.



(84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE,

SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

SYSTEM AND METHOD FOR DYNAMIC RECOVERY OF AUDIO
DATA AND COMPRESSED AUDIO ENHANCEMENT

TECHNICAL FIELD

[0001] The present disclosure relates generally to audio processing, and more specifically to a system and method for compressed audio enhancement that improves the perceived quality of the compressed audio signal to the listener.

BACKGROUND OF THE INVENTION

[0002] Compressed audio data is notoriously poor in quality. Despite this problem, no known solutions exist to improve the listener experience.

SUMMARY OF THE INVENTION

[0003] A system for processing digitally-encoded audio data is provided that includes a compressed audio source device providing a sequence of frames of compressed digital audio data. A compressed audio enhancement system receives the sequence of frames of compressed digital audio data and generates enhanced audio data by adding masked digital audio data to the sequence of frames of compressed digital audio data, where the masked digital audio data has an energy level sufficient to keep a kinocilia of a listener active. One or more speakers configured to receive the enhanced audio data and to generate sound waves using the enhanced audio data.

[0004] Other systems, methods, features, and advantages of the present disclosure will be or become apparent to one with skill in the art upon examination of the following drawings and detailed description. It is intended that all such additional systems, methods, features, and advantages be included within

this description, be within the scope of the present disclosure, and be protected by the accompanying claims.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0005] Aspects of the disclosure can be better understood with reference to the following drawings. The components in the drawings may be, but are not necessarily, to scale, emphasis instead being placed upon clearly illustrating the principles of the present disclosure. Moreover, in the drawings, like reference numerals designate corresponding parts throughout the several views, and in which:

[0006] FIGURE 1 is a diagram of a frequency diagram showing the effect of compressed audio processing in accordance with the present disclosure;

[0007] FIGURE 2 is a diagram of a system for enhancing compressed audio data with controlled modulation distortion in accordance with an exemplary embodiment of the present disclosure;

[0008] FIGURE 3 is a diagram of a system for providing modulation distortion in accordance with an exemplary embodiment of the present disclosure;

[0009] FIGURE 4 is a diagram of an algorithm for processing compressed audio data to provide kinocilia stimulation, in accordance with an exemplary embodiment of the present disclosure;

[0010] FIGURE 5 is a diagram of a system for processing audio data to provide for dynamic recovery of audio data, in accordance with an exemplary embodiment of the present disclosure;

[0011] FIGURE 6 is a diagram of an algorithm for processing audio data to provide for dynamic recovery of audio data, in

accordance with an exemplary embodiment of the present disclosure;

[0012] FIGURE 7 is a diagram of a system for providing dynamic volume control in accordance with an exemplary embodiment of the present disclosure;

[0013] FIGURE 8 is a diagram of a system for television volume control, in accordance with an exemplary embodiment of the present disclosure;

[0014] FIGURE 9 is a diagram of an algorithm for processing television audio data to provide volume control, in accordance with an exemplary embodiment of the present disclosure;

[0015] FIGURE 10 is a diagram of a system for decorrelating audio in accordance with an exemplary embodiment of the present disclosure;

[0016] FIGURE 11 is a diagram of an array of randomly variable delays in accordance with an exemplary embodiment of the present disclosure;

[0017] FIGURE 12 is a diagram of an algorithm in accordance with an exemplary embodiment of the present disclosure;

[0018] FIGURE 13 is a diagram of a system for dynamic equalization of audio data, in accordance with an exemplary embodiment of the present disclosure;

[0019] FIGURE 14 is a diagram of a system for controlling dynamic equalization of audio data, in accordance with an exemplary embodiment of the present disclosure;

[0020] FIGURE 15 is a diagram of an algorithm for dynamic equalization of audio data, in accordance with an exemplary embodiment of the present disclosure;

[0021] FIGURE 16 is a diagram of a system for parametric stereo processing, in accordance with an exemplary embodiment of the present disclosure;

[0022] FIGURE 17 is a diagram of an algorithm for parametric stereo processing, in accordance with an exemplary embodiment of the present disclosure;

[0023] FIGURE 18 is a diagram of a system for providing artifact masking of audio data artifacts in accordance with an exemplary embodiment of the present disclosure; and

[0024] FIGURE 19 is a diagram of an algorithm for processing audio data in accordance with an exemplary embodiment of the present disclosure.

DETAILED DESCRIPTION OF THE INVENTION

[0025] In the description that follows, like parts are marked throughout the specification and drawings with the same reference numerals. The drawing figures may be to scale and certain components can be shown in generalized or schematic form and identified by commercial designations in the interest of clarity and conciseness.

[0026] FIGURE 1 is a diagram of a frequency diagram 100 showing the effect of compressed audio processing to create double sideband AM components in audio signals, in accordance with the present disclosure. Frequency diagram 100 shows a frequency distribution for audio data, such as compressed audio data, with frequency components at $\pm F1$, $F2$ and $F3$. These frequency components are relatively sparse, such as may result from audio data that has been compressed using lossy compression techniques. Frequency diagram 100 also shows a frequency distribution for enhanced compressed audio data with dynamic recovery, having frequency components centered at $\pm F1$, $F2$ and $F3$ and associated modulation distortion components (such as double sideband AM components) in a range around the centered frequency components. Although the modulation distortion components/double sideband AM components are represented as

having an essentially flat profile, a Gaussian distribution, an exponential decay or any suitable profile can also or alternatively be used. The magnitude of the modulation distortion components/double sideband AM components is also at least 13 dB below the signal magnitude, in order to mask the modulation distortion components from perception by the user.

[0027] Typically, modulation distortion, such as double sideband AM components, is avoided. However, the present disclosure recognizes that kinocilia require a certain level of stimulation to remain in an active state, and otherwise will go into a dormant state, until a threshold level of audio energy causes them to switch from the dormant state to the active state. By generating modulation distortion/double sideband AM components, the kinocilia can be stimulated to remain in the active state, even if the audio signals are masked by being more than 13 dB in magnitude relative to a major frequency component. The use of modulation distortion such as double sideband AM components in this manner enhances the audio listening experience, because the kinocilia remain active and can detect frequency components of the compressed audio data that would otherwise not have sufficient energy to switch them out of the dormant state.

[0028] **FIGURE 2** is a diagram of a system 200 for enhancing compressed audio data with controlled modulation distortion in accordance with an exemplary embodiment of the present disclosure. System 200 includes compressed audio source device 202, compressed audio enhancement 204 and speakers 206, each of which are specialized devices or apparatuses that can be implemented in hardware or a suitable combination of hardware and software.

[0029] As used herein, the singular forms "a", "an" and "the" are intended to include the plural forms as well, unless the

context clearly indicates otherwise. It will be further understood that the terms "comprises" and/or "comprising," when used in this specification, specify the presence of stated features, integers, steps, operations, elements, and/or components, but do not preclude the presence or addition of one or more other features, integers, steps, operations, elements, components, and/or groups thereof. As used herein, the term "and/or" includes any and all combinations of one or more of the associated listed items. As used herein, phrases such as "between X and Y" and "between about X and Y" should be interpreted to include X and Y. As used herein, phrases such as "between about X and Y" mean "between about X and about Y." As used herein, phrases such as "from about X to Y" mean "from about X to about Y."

[0030] As used herein, "hardware" can include a combination of discrete components, an integrated circuit, an application-specific integrated circuit, a field programmable gate array, or other suitable hardware. As used herein, "software" can include one or more objects, agents, threads, lines of code, subroutines, separate software applications, two or more lines of code or other suitable software structures operating in two or more software applications, on one or more processors (where a processor includes one or more microcomputers or other suitable data processing units, memory devices, input-output devices, displays, data input devices such as a keyboard or a mouse, peripherals such as printers and speakers, associated drivers, control cards, power sources, network devices, docking station devices, or other suitable devices operating under control of software systems in conjunction with the processor or other devices), or other suitable software structures. In one exemplary embodiment, software can include one or more lines of code or other suitable software structures operating in a

general purpose software application, such as an operating system, and one or more lines of code or other suitable software structures operating in a specific purpose software application. As used herein, the term "couple" and its cognate terms, such as "couples" and "coupled," can include a physical connection (such as a copper conductor), a virtual connection (such as through randomly assigned memory locations of a data memory device), a logical connection (such as through logical gates of a semiconducting device), other suitable connections, or a suitable combination of such connections.

[0031] Compressed audio source device 202 provides a stream of digitally-encoded audio data, such as frames of encoded digital data, from a memory storage device such as a random access memory that has been configured to store digitally-encoded audio data, from an optical data storage medium that has been configured to store digitally-encoded audio data, from a network connection that has been configured to provide digitally-encoded audio data, or in other suitable manners. Compressed audio source device 302 can be implemented as a special purpose device such as an audio music player, a cellular telephone, an automobile audio system or other suitable audio systems that are configured to provide streaming compressed audio data.

[0032] Compressed audio enhancement 204 is coupled to compressed audio source device 202, such as by using a wireless or wireline data communications medium. Compressed audio enhancement 204 enhances the compressed audio data for a listener by introducing modulation distortion or by otherwise introducing audio signal components that are masked by the compressed audio signal data but which are of sufficient magnitude to stimulate the kinocilia of the listener, so as to prevent the kinocilia from switching to a dormant state that

requires a substantially higher amount of energy to switch back to an active state than might be provided by the compressed audio data at any given instant. By keeping the kinocilia in an active state, compressed audio enhancement 204 improves the ability of the listener to hear the audio signals encoded in the compressed audio data.

[0033] Speakers 206 receive the enhanced compressed audio data and generate sound waves that can be perceived by a listener. Speakers 206 can be implemented as mono speakers, stereo speakers, N.1 surround speakers, automobile speakers, headphone speakers, cellular telephone speakers, sound bar speakers, computer speakers or other suitable speakers.

[0034] In operation, system 200 enhances compressed and digitally-encoded audio data by introducing additional frequency components that are masked by the compressed and digitally-encoded audio data, but which are of a sufficient magnitude to keep the listener's kinocilia active. In this manner, the listener is able to hear additional compressed and digitally-encoded audio data signals that would otherwise not be perceived, which results in an improved listening experience.

[0035] **FIGURE 3** is a diagram of a system 300 for providing modulation distortion in accordance with an exemplary embodiment of the present disclosure. System 300 includes high pass filter 302, low pass filter 304, Hilbert transform 306, summation unit 308, high pass filter 310 and modulation distortion 312, each of which can be implemented in hardware or a suitable combination of hardware and software.

[0036] High pass filter 302 is configured to receive compressed and digitally-encoded audio data signals and to filter out low frequency components from the signal. In one exemplary embodiment, high pass filter 302 can be implemented as a high-pass order 1 Butterworth filter having a 118 Hz corner

frequency, using the Audio Weaver design environment from DSP Concepts or other suitable design environments, hardware or hardware and software.

[0037] Low pass filter 304 is coupled to high pass filter 302 and is configured to receive the filtered, compressed and digitally-encoded audio data signals and to filter out high frequency components from the signal. In one exemplary embodiment, low pass filter 304 can be implemented as a low-pass order 4 Butterworth filter having a 10400 Hz corner frequency, using the Audio Weaver design environment from DSP Concepts or other suitable design environments, hardware or hardware and software.

[0038] Hilbert transform 306 is coupled to low pass filter 304 and is configured to receive the filtered, compressed and digitally-encoded audio data signals and to apply a Hilbert transform to the signal. In one exemplary embodiment, Hilbert transform 306 can be implemented using the Audio Weaver design environment from DSP Concepts or other suitable design environments, hardware or hardware and software, and can receive a split output from low pass filter 304 and can apply a Hilbert +/- 90 degree phase shift to each output.

[0039] Summation unit 308 is coupled to Hilbert transform 306 and is configured to square each split output signal and to then take the square root of the sum, in order to obtain an absolute value of the signal.

[0040] High pass filter 310 is coupled to summation unit 308 and is configured to filter out low frequency components from the signal. In one exemplary embodiment, high pass filter 310 can be implemented as a high-pass order 2 Butterworth filter having a 1006 Hz corner frequency, using the Audio Weaver design environment from DSP Concepts or other suitable design environments, hardware or hardware and software.

[0041] Modulation distortion 312 is coupled to high pass filter 310 and is configured to receive the filtered and compressed audio signal data and to add modulation distortion to the data. In one exemplary embodiment, modulation distortion 312 can be implemented using a downward expander of the Audio Weaver design environment from DSP Concepts or other suitable design environments, hardware or hardware and software. In this exemplary embodiment, the downward expander can be implemented as a software system operating on a specialized processor that has a plurality of settings, such as a threshold setting, a ratio setting, a knee depth setting, an attack time setting, a decay time setting and other suitable settings. These settings can be selected to optimize the generation of modulation distortion in the vicinity of the frequency components of the compressed audio data, such as by setting the threshold setting to a range of -23 dB +/- 20%, the ratio setting to range of 1.0 to 1.016 dB/dB +/- 20%, the knee depth setting to 0 dB +/- 20%, the attack time setting to 0.01 milliseconds +/- 20%, the decay time setting to 3 milliseconds +/- 20% or in other suitable manners. In general, the attack time may have the greatest influence on generation of phase distortion, and a setting of 1 millisecond or less can be preferable. These exemplary settings can result in the generation of modulation distortion, which is typically avoided, but which is used in this exemplary embodiment specifically to cause the compressed and digitally-encoded audio data to have modulation distortion signals that are below a perceptual threshold by virtue of being masked by the encoded signal components. As the encoded signal components change over time, the kinocilia in the frequency range surrounding the encoded signal components are stimulated enough to prevent them from switching from an active state to a dormant state, thus ensuring that they are able to detect encoded audio

signals that are at a magnitude that would otherwise be insufficient to cause dormant kinocilia to switch to an active state. The output of modulation distortion 312 can be provided to an amplifier, a speaker or other suitable devices.

[0042] In operation, system 300 provides optimal audio signal processing for compressed audio data to provide a level of modulation distortion that is below a perceptual level but which is sufficient to improve the quality of the listening experience, by providing sufficient stimulation to the kinocilia to prevent them from switching from an active to a dormant state. In this manner, the listening experience is improved, because the listener can perceive audio signals that would otherwise not be perceived.

[0043] FIGURE 4 is a diagram of an algorithm 400 for processing compressed audio data to provide kinocilia stimulation, in accordance with an exemplary embodiment of the present disclosure. Algorithm 400 can be implemented in hardware or a suitable combination of hardware and software, and can be one or more software systems operating on a special purpose processor.

[0044] Algorithm 400 begins at 402, where compressed audio data is received from a source device. In one exemplary embodiment, a frame of the compressed audio data can be received at an input port to an audio data processing system and stored to a buffer device, such as random access memory that has been configured to store audio data. In addition, a processor can be configured to sense the presence of the audio data, such as by checking a flag or other suitable mechanism that is used to indicate that audio data is available for processing. The algorithm then proceeds to 404.

[0045] At 404, low frequency components are removed from the audio data. In one exemplary embodiment, a high pass filter can

be used to filter out low frequency components from the audio data, such as a high-pass order 1 Butterworth filter having a 118 Hz corner frequency or other suitable filters. The filtered audio data can then be stored in a new buffer, in the same buffer or in other suitable manners. The algorithm then proceeds to 406.

[0046] At 406, high frequency components are removed from the audio data. In one exemplary embodiment, a low pass filter can be used to filter out high frequency components from the signal, such as a low-pass order 4 Butterworth filter having a 10400 Hz corner frequency or other suitable filters. The filtered audio data can then be stored in a new buffer, in the same buffer or in other suitable manners. The algorithm then proceeds to 408.

[0047] At 408, Hilbert filtering is performed on the low pass filtered data, to generate two sets of data having a +/- 90 degree phase shift. The Hilbert filtered audio data can then be stored in a new buffer, in the same buffer or in other suitable manners. The algorithm then proceeds to 410.

[0048] At 410, the absolute value of the signal is obtained, such as by using a summation unit that is configured to square each set of data and to then take the square root of the sum, in order to obtain an absolute value of the signal, or in other suitable manners. The absolute value audio data can then be stored in a new buffer, in the same buffer or in other suitable manners. The algorithm then proceeds to 412.

[0049] At 412, the absolute value data is filtered to remove low frequency components from the signal, such as by using a high-pass order 2 Butterworth filter having a 1006 Hz corner frequency or in other suitable manners. The filtered audio data can then be stored in a new buffer, in the same buffer or in other suitable manners. The algorithm then proceeds to 414.

[0050] At 414, modulation distortion is generated in the audio data, such as by processing the audio data using a downward expander having a threshold setting of -23 dB, a ratio setting of 1.016 dB/dB, a knee depth setting of 0 dB, an attack time setting of 0.01 milliseconds, a decay time setting of 3 milliseconds or other suitable settings. These exemplary settings can result in the generation of modulation distortion, which is typically avoided, but which is used in this exemplary embodiment specifically to cause the compressed and digitally-encoded audio data to have modulation distortion signals that are below a perceptual threshold by virtue of being masked by the encoded signal components. As the encoded signal components change over time, the kinocilia in the frequency range surrounding the encoded signal components are stimulated enough to prevent them from switching from an active state to a dormant state, thus ensuring that they are able to detect encoded audio signals that are at a magnitude that would otherwise be insufficient to cause dormant kinocilia to switch to an active state. The algorithm then proceeds to 416.

[0051] At 416, the processed audio data is output to an amplifier, a speaker or other suitable devices. In one exemplary embodiment, the processed audio data can be stored in a buffer and can be retrieved periodically for provision to a digital signal processor, a digital to analog converter, an amplifier or other suitable devices for generation of an analog signal that is provided to speakers.

[0052] **FIGURE 5** is a diagram of a system 500 for processing audio data to provide for dynamic recovery of audio data, in accordance with an exemplary embodiment of the present disclosure. System 500 can be implemented in hardware or a suitable combination of hardware and software.

[0053] System 500 includes AGC core 502, which receives an audio input signal INPUT and which processes the audio to remove a DC signal component, such as to maintain the audio input signal at an average level. In one exemplary embodiment, the audio input signal can be averaged over 200 mS period or other suitable periods.

[0054] AGC multiplier 504 receives the unprocessed audio input signal and the output of AGC core 502 and generates a normalized audio output signal.

[0055] High pass filter 506 and low pass filter 508 are used to form a band pass filter over a range that is generally too large to accommodate a single band pass filter element, such as with a lower frequency cutoff of about 100 to 130 Hz, for example 118 Hz, and a higher frequency cutoff of 10,000 to 11,000 Hz, for example 10400 Hz, or other suitable frequency ranges.

[0056] Hilbert converter 510 converts the filtered audio data to Hilbert space, by shifting one channel input 90 degrees from the other, to form a Bedrosian all pass pair. When signals are at 90 degrees to each other, they are orthogonal and decorrelated. The output of Hilbert converter 510 is provided to square add 512, to yield an instantaneous magnitude envelope detector. Scaler 514 is used to provide gain control to the output of square add 512.

[0057] Butterworth filter 516 is a second order high pass filter that is used to remove low frequency components that do not have as much of an effect on audio quality. For example, for frequency components below 1200 Hz, only the phase is significant for the "cocktail party" effect and other hearing processes that are used to improve audio quality, whereas above 1200 Hz, the envelope and magnitude of the signal is more

important. Maximum absolute value detector 518 is used to generate the maximum absolute value of the signal.

[0058] Downward expander 520 is used to generate a double sideband suppressed AM signal. The release time (also shown as decay time or persistence) changes spread of side bands, and a suitable release time can be used, such as 3 milliseconds or less, although a suitable setting in the range of 1 millisecond to 8 milliseconds can be used. A short attack can be used, such as 0.01 milliseconds. Increasing the signal to noise ratio can be used to increase the height of the side band, where a setting of 1.06 to 1.09 can be used. The release time (decay time/persistence) changes spread of side bands.

[0059] The double sideband suppressed AM signal is used to stimulate kinociliar recruitment, which can be used when there is a hole or frequency gap in the frequency response of the Cochlea. By stimulating the kinocilia on either side of frequency gap, the frequency response of the damaged kinocilia can be simulated to the listener. Furthermore, when the sideband signals are present at 13 dB below the primary audio signal, the simulated signal is not audible over the primary audio signal.

[0060] The use of downward expander 520 in this manner also helps to improve the audio quality of audio that is generated using lossy compression, which increases spectrum sparsity. When there are less frequency components in content, the image width narrows, and the kinocilia are also under-stimulated, which can result in the active kinocilia becoming dormant. When that occurs, additional audio energy must be provided to activate the dormant kinocilia.

[0061] AGC multiplier 524 is used to combine the double sideband suppressed AM signal with the original audio signal, which is delayed by delay 522 to equal the time required for

processing in the upper chain. Scaler 526 is used to provide gain control to the processed audio data.

[0062] In operation, system 500 provides dynamic recovery of audio data to improve the audio quality of audio data that is generated from stored or transmitted audio data that resulted from lossy compression processes. The dynamic recovery of audio data generates a double sideband suppressed AM signal having a magnitude that is 13 dB lower than the associated primary audio signal components, which can both improve the image width and audio quality of the audio data as well as simulate the stimulation of missing frequencies of kinocilia for listeners with hearing loss. In this manner, system 500 can not only improve audio quality for persons without hearing loss, but can also improve audio quality for persons with such hearing loss.

[0063] **FIGURE 6** is a diagram of an algorithm 600 for processing audio data to provide for dynamic recovery of audio data, in accordance with an exemplary embodiment of the present disclosure. Algorithm 600 can be implemented in hardware or a suitable combination of hardware and software.

[0064] Algorithm 600 begins at 602, where an input audio signal is processed to remove DC signal components, such as by averaging the signal over 200 milliseconds and then processing with an automatic gain control processor, or in other suitable manners. The algorithm then proceeds to 604.

[0065] At 604, the audio signal is level normalized, such as by multiplying the DC-compensated audio data with the unprocessed audio data in an automatic gain control processor or in other suitable manners. The algorithm then proceeds to 606.

[0066] At 606, the normalized audio signal is band pass filtered, such as by using a separate low pass filter and high pass filter or in other suitable manners. The algorithm then proceeds to 608.

[0067] At 608, the audio signal is processed using a Hilbert converter, such as to shift a left channel signal relative to a right channel signal by 90 degrees, to de-correlate the channels of audio data, or in other suitable manners. The algorithm then proceeds to 610.

[0068] At 610, the outputs from the Hilbert converter are squared and added, such as to generate an instantaneous magnitude envelope detector or for other suitable purposes. The algorithm then proceeds to 612.

[0069] At 612, the gain of the envelop signal is adjusted. The algorithm then proceeds to 614.

[0070] At 614, the processed audio signal is filtered using a high pass filter to remove audio components that are not susceptible to improvement using dynamic recovery, such as audio signal components having a frequency below 1000 to 1200 Hz. The algorithm then proceeds to 616.

[0071] At 616, the absolute value of the audio signal is generated. The algorithm then proceeds to 618.

[0072] At 618, the absolute value is processed using a downward expander to generate a double sideband suppressed AM signal having a magnitude that is 13 dB lower than the associated primary audio signal components. The algorithm then proceeds to 620.

[0073] At 620, the double sideband suppressed AM signal is mixed with a delayed corresponding input audio signal to generate audio data with improved quality. The algorithm then proceeds to 622.

[0074] At 622, the gain of the audio data is adjusted, and the processed audio is output.

[0075] In operation, algorithm 600 is used to generate audio data that includes dynamically-recover audio components, to improve the quality of the audio data. In one exemplary

embodiment, algorithm 600 can be used to improve audio quality for audio data that is generated from lossy-compression output (where the number of frequency components in the audio data is lower than normal), by generating double sideband suppressed AM signal components adjacent to the primary frequency components, but where the double sideband suppressed AM signal has a magnitude of at least 13 dB lower than the primary signal. In this manner, the double sideband suppressed AM signals are not audible to the listener, but provide sufficient energy to the kinocilia to keep them stimulated, so as to reduce the amount of energy required to activate them. In another exemplary embodiment, the double sideband suppressed AM signal can aid a listener with hearing loss resulting from loss of kinocilia by simulating the stimulation of the missing frequency bands through stimulation of the adjacent frequency components.

[0076] Level control for television programming typically requires processing at the head end, the affiliate and at the receiver or television set. The lack of consistent coordination between these three entities has resulted in a lack of effective volume control, which results in large swings between content that is too quiet and content that is too loud.

[0077] **FIGURE 7** is a diagram of a system 700 for providing dynamic volume control in accordance with an exemplary embodiment of the present disclosure. System 700 can be implemented in hardware or a suitable combination of hardware and software.

[0078] System 700 includes television 702, which includes television signal processor 704, dynamic volume control 706 and speaker 708. Television signal processor 704 receives an encoded television signal and processes the signal to extract audio and video signals. The video signals are provided to a video display, and the audio signals are generally provided to an

amplifier or other audio systems. However, television audio signals have inconsistent volume settings because they originate from different sources, such as programming, national commercial content and local commercial content. As a result, the user experience is often unsatisfactory, because the sound can go from a user setting to an extreme loud setting or an extreme quiet setting without user action.

[0079] To address these problems, dynamic volume control 706 can detect a programming type and a time of day, and can apply dynamic volume processing to the television audio signal. In one exemplary embodiment, dynamic volume control 706 can use a series of automatic gain control (AGC) units that are dynamically increased or decreased as a function of the time of day and the type of programming, and can adjust the settings of each AGC accordingly. The use of a series of AGC units with the proper settings provides optimal volume control, which effectively increases or decreases sound to a preferred level without resulting in overshoots, undershoots or other undesired audio artifacts. The audio signals that have been processed by dynamic volume control 706 to provide volume control are then provided to speaker 708, which generates sound waves having the desired volume control qualities.

[0080] **FIGURE 8** is a diagram of a system 800 for television volume control, in accordance with an exemplary embodiment of the present disclosure. System 800 includes AGC units 802A through 802N, time of day system 804 and programming type system 806. In one exemplary embodiment, the number of AGC units and the settings for the AGC units can be varied by time of day system 804 as a function of the time of day and by programming type system 806 as a function of the type of programming.

[0081] In this exemplary embodiment, the first AGC unit such as AGC unit 802A can have a target level setting in dB that is

the lowest of the series of AGC units, such as -34 dB, and can have an associated activation threshold that is also the lowest of the series of AGC units, such as -40 dB. The first AG unit 802A can also have a longest smoothing time, such as 800 milliseconds. The last AGC unit 802N can have a target level setting that is the highest of the series of AGC units, such as -24 dB, and the associated activation threshold can also be the highest, such as -30dB. In this exemplary embodiment, the difference between the activation threshold and the target level can be set to the same value, such as 6 dB, and the activation threshold and target values can be increased as a function of the number of AGC units. For example, for 6 AGC units, the activation levels can increase from -40 dB to -38 dB, -36 dB, -34 dB, -32 dB and -30 dB, and the target values can increase from -34 dB to -32 dB, -30 dB, -28 dB, -26 dB and -24 dB. Likewise, the smoothing time can decrease from 200 milliseconds (mS) to 100 mS, 50 mS, 25 mS, 10 mS and 1 mS. The recovery rate for all AGC units can be set to a low value, such as 1 dB/second.

[0082] The use of a series of AGC units with the disclosed settings helps to reduce rapid increases in volume that are of a short duration, while also decreasing longer-term increases in volume more slowly and over a longer period of time. In this manner, control of variations in volume caused by different programing sources can be controlled at the television set, regardless of whether any controls are implemented by the broadcaster or local affiliate.

[0083] **FIGURE 9** is a diagram of an algorithm 900 for processing television audio data to provide volume control, in accordance with an exemplary embodiment of the present disclosure. Algorithm 900 can be implemented in hardware or a

suitable combination of hardware and software, and can be one or more software systems operating on a special purpose processor.

[0084] Algorithm 900 begins at 902, where the time of day is detected. In one exemplary embodiment, the time of day can be obtained from a network signal, a local signal, a local timekeeping device or other suitable sources, such as by implementing one or more controls that cause a television control processor to query a data buffer, to transmit a data query to a predetermined address or in other suitable manners. The algorithm then proceeds to 904.

[0085] At 904, a content type is detected. In one exemplary embodiment, a content type indicator can be obtained from a network signal, a local signal, a local directory that lists programs on each channel or other suitable sources, such as by implementing one or more controls that cause a television control processor to query a data buffer, to transmit a data query to a predetermined address or in other suitable manners. The algorithm then proceeds to 906.

[0086] At 906, a number of AGC units is selected for a type of content, a time of day or for other suitable program parameters. In one exemplary embodiment, the number of AGC units can be determined from a look-up table, based on range values associated with a program, based on historical data stored at a television set, from a network signal, from a local signal or from other suitable sources, such as by implementing one or more controls that cause a television control processor to query a data buffer, to transmit a data query to a predetermined address or in other suitable manners. The algorithm then proceeds to 908.

[0087] At 908, the target values for the AGC units are set from low to high. In one exemplary embodiment, a television control processor can determine and store a target value for

each AGC unit as a function of the number of AGC units, target values can be retrieved from memory or other suitable processes can also or alternatively be used. In this exemplary embodiment, beginning and ending target values for a time of day or type of program can be determined, such as from a data table, and steps between the beginning and ending target values can be generated as a function of the number of AGC units, such as by applying a linear transition scheme (e.g. -34 dB, -32 dB, -30 dB, -28 dB, -26 dB and -24 dB), a nonlinear transition scheme (e.g. -34 dB, -31 dB, -29 dB, -28 dB, -27 dB and -24 dB) or in other suitable manners. The algorithm then proceeds to 910.

[0088] At 910, the activation threshold levels for the AGC units are set from low to high. In one exemplary embodiment, a television control processor can determine and store an activation threshold value for each AGC unit as a function of the number of AGC units, activation threshold values can be retrieved from memory or other suitable processes can also or alternatively be used. In this exemplary embodiment, beginning and ending activation threshold values for a time of day or type of program can be determined, such as from a data table, and steps between the beginning and ending activation threshold values can be generated as a function of the number of AGC units, such as by applying a linear transition scheme (e.g. -40 dB, -38 dB, -36 dB, -32 dB, -30 dB and -28 dB), a nonlinear transition scheme (e.g. -40 dB, -37 dB, -35 dB, -34 dB, -33 dB and -28 dB) or in other suitable manners. The algorithm then proceeds to 912.

[0089] At 912, the smoothing or activation time values for the AGC units are set from slow to fast. In one exemplary embodiment, a television control processor can determine and store a smoothing time value for each AGC unit as a function of the number of AGC units, smoothing time values can be retrieved

from memory or other suitable processes can also or alternatively be used. In this exemplary embodiment, beginning and ending smoothing time values for a time of day or type of program can be determined, such as from a data table, and steps between the beginning and ending smoothing time values can be generated as a function of the number of AGC units, such as by applying a linear transition scheme (e.g. 200 mS, 150 mS, 100 mS, 50 mS and 0 mS), a nonlinear transition scheme (e.g. 200 mS, 100 mS, 50 mS, 25 mS and 0 mS) or in other suitable manners. The algorithm then proceeds to 914.

[0090] At 914, the recovery rate values for the AGC units are set. In one exemplary embodiment, a television control processor can determine and store a recovery rate value for each AGC unit as a function of the number of AGC units, recovery rate values can be retrieved from memory or other suitable processes can also or alternatively be used. The recovery rate for each AGC unit can be set to the same value, such as 1 dB/sec., or other suitable settings can also or alternatively be used. The algorithm then proceeds to 916.

[0091] At 916, the audio data is processed. In one exemplary embodiment, the audio data can be extracted from encoded television programming signals and processed through the series of AGC units to control a volume of the audio data. The algorithm then proceeds to 918.

[0092] At 918, it is determined whether a content type has changed. If it is determined that a content type has changed, the algorithm returns to 904, otherwise the algorithm proceeds to 920.

[0093] At 920, it is determined whether a time of day has changed. If it is determined that a time of day has changed, the algorithm returns to 902, otherwise the algorithm returns to 316.

[0094] In operation, algorithm 900 allows volume settings of television programming audio data to be automatically adjusted based on the time of day, the type of program and other suitable factors, in order to automatically adjust the volume to compensate for sudden volume changes caused by programming at the national level, at the local level or in other suitable manners.

[0095] FIGURE 10 is a diagram of a system 1000 for decorrelating audio in accordance with an exemplary embodiment of the present disclosure. System 1000 can be implemented in hardware or a suitable combination of hardware and software, and can be one or more software systems operating on a special purpose audio processor or other suitable devices.

[0096] System 1000 includes adder 1002, which receives an input audio data stream and adds it to a feedback audio data stream. The output of adder 1002 is provided to router 1004, which routes the audio data to the inputs of scattering matrix 1006. Scattering matrix 1006 can utilize a full M-input x N-output mixer with a total of M x N coefficients. The coefficients can be contained in a coefficients matrix with the gain from the *i*th input to the *j*th output found at `coeffs(i, j)`. In this exemplary embodiment, the `coeffs` array can be a one dimensional array of length M x N. The gain from input *m* ($0 \leq m \leq M-1$) to output *n* ($0 \leq n \leq N-1$) can be stored at `coeffs[m + n*M]`. In this manner, the first M coefficients can be the gains used to compute the first output channel. The next M coefficients can be the gains used to compute the second output channel, and so forth.

[0097] The output from scattering matrix 1006 is provided to delay 1008, which is a randomly variable delay. By introducing a randomly variable delay with less than 5 cents of variation, the audio data is decorrelated from other audio data as well as

itself, such as to prevent the audio data from exciting resonance modes in chassis or other structures. The output of delay 1008 is provided to low pass filter 1010, which is then provided to scaler 1012.

[0098] Scaler 1012 is used to reduce the level of the decorrelated audio data to at least 13 dB below the unprocessed audio data, so that it is present in the audio data but not noticeable to a human listener. This effect is based on the observation that an audio content signal that is 13 dB or greater than associated an audio noise signal will effectively block the audio noise signal from being perceived by a listener. The output from scaler 1012 is fed back to the input of router 1004, as well as to router 1014 and router 1016. The output of router 1014 is fed back to the input to adder 1002, and the output of router 116 is fed to mixer 1018, where it is mixed with the audio input signal.

[0099] In operation, system 1000 receives and processes input digital audio data in the time domain to decorrelate the digital audio data, such as to combine overlay audio data that varies slightly from the input audio data with the input audio data to generate decorrelated output audio data. The decorrelated audio data is then combined with the unprocessed audio data at a level that is at least 13 decibels below the level of the unprocessed audio data, so as to cause the kinocilia to be stimulated at a sound level that is not perceived by the listener, but which nonetheless prevents the listener's kinocilia from becoming dormant, which would require additional audio energy to wake the kinocilia and which effectively masks low level audio data in the unprocessed audio data. Although time domain processing is disclosed, frequency domain components and processing can also or alternatively be used where suitable.

[00100] **FIGURE 11** is a diagram of an array 1100 of randomly variable delays in accordance with an exemplary embodiment of the present disclosure. Array 1100 receives the output of scattering matrix 1006 into each of a plurality of delay units 1104A through 1104N, which each have an associated random input 1102A through 1102N. In one exemplary embodiment, each delay unit can be a Z^{-N} delay unit, where the value of N is changed by an amount not exceeding 5 cents. The outputs of delays 1104A through 1104N are then summed at summer 1106 to generate a decorrelated output signal.

[00101] Although a plurality of programmable delays 1104A through 1104N are shown with a plurality of random inputs, a single programmable delay with a single random input can be used, as long as the delay is truly randomized. When a non-random delay is used, patterns in the delay variation can be detectable to human hearing, or can otherwise excite resonant oscillation modes in speaker grills, enclosures or other structures.

[00102] **FIGURE 12** is a diagram of an algorithm 1200 in accordance with an exemplary embodiment of the present disclosure. Algorithm 1200 can be implemented in hardware or a suitable combination of hardware and software.

[00103] Algorithm 1200 begins at 1200, where audio data is received. In one exemplary embodiment, the audio data can be a single channel of audio data in a digital format that is received over a data port and buffered in a data memory, multiple channels of audio data in a digital format that are received in data packets having a multiple-channel format and which are processed to extract individual channels that are buffered into separate data memory locations, one or multiple channels of audio data in a predetermined analog format that are subsequently converted to a digital format using an analog to

digital data processing system and buffered in a data memory, or other suitable forms of audio data. The algorithm then proceeds to 1204.

[00104] At 1204, the audio data is mixed with audio data that has been fed back after processing. In one exemplary embodiment, the mixing can be performed by simple addition of the two audio signals, or other suitable mixing processes can also or alternatively be used. The algorithm then proceeds to 1206.

[00105] At 1206, the mixed audio data is routed to a scattering matrix. In one exemplary embodiment, the audio data can be routed as different time samples, as redundant data streams or in other suitable manners. The algorithm then proceeds to 1208.

[00106] At 1208, the audio data is processed by the scattering matrix, such as in the manner discussed herein or in other suitable manners. The output of the scattering matrix can then be summed or reconstructed as suitable, and the algorithm proceeds to 1210.

[00107] At 1210, a variable delay is applied to the audio data. In one exemplary embodiment, the audio data can be processed using a Z^{-N} transform, where the delay value is changed by an amount not exceeding 5 cents (.416 percent). In one exemplary embodiment, the amount of delay can be limited to no more than an amount that is below a level at which human hearing can distinguish a change in a steady single frequency tone, so as to prevent the random variation from being noticeable. In one exemplary embodiment, the output of the scattering matrix can be provided in parallel to a plurality of parallel programmable delays, which each have a distinct random variation to the amount of delay, and where the audio output data from each delay is then added together to form a single set of audio data, or

other suitable processing can also or alternatively be used. The algorithm then proceeds to 1212.

[00108] At 1212, the audio data is processed by a low pass filter. In one exemplary embodiment, the low pass filter can have an 8 kilohertz cutoff frequency or other suitable frequencies, such as to extract higher frequency components where pitch variations are difficult to hear. The algorithm then proceeds to 1214.

[00109] At 1214 the audio data is processed by a scaler, such as to reduce the magnitude of the audio data prior to feeding the audio data back for processing at 1202 and 1204, for mixing with input audio data or for other suitable purposes. In one exemplary embodiment, the level can be sufficient to prevent kinocilia from becoming inactive for more than 200 milliseconds, because the amount of energy required to re-active kinocilia after they have become inactive can result in a reduction in the perceived quality of the audio data.

[00110] At 1216, the audio data is mixed with the input audio data and is fed back for mixing at 1202 and is also fed back for routing at 1204. As discussed, the use of decorrelated audio that has been added into the original audio data can improve the sensitivity of kinocilia to the processed audio data by preventing the kinocilia from becoming dormant.

[00111] In one exemplary embodiment, the present disclosure contemplates a system for processing digitally-encoded audio data that has a compressed audio source device providing a sequence of frames of compressed digital audio data, a compressed audio enhancement system configured to receive the sequence of frames of compressed digital audio data and to generate enhanced audio data by adding masked digital audio data to the sequence of frames of compressed digital audio data, where the masked digital audio data has an energy level

sufficient to keep a kinocilia of a listener active, one or more speakers configured to receive the enhanced audio data and to generate sound waves using the enhanced audio data, a high pass filter configured to remove low frequency components of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, a low pass filter configured to remove high frequency components of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, a Hilbert transform configured to apply a phase shift to the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, an absolute value processor configured to generate an absolute value of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, wherein generation of the enhanced audio data comprises generating modulation distortion of the enhanced audio data, wherein generation of the enhanced audio data comprises generating modulation distortion for one or more frequency components of the enhanced audio data, wherein generation of the enhanced audio data comprises generating modulation distortion for one or more frequency components of the enhanced audio data, the modulation distortion having a frequency range centered at each of the associated frequency components, wherein the modulation distortion has a magnitude at least 13 dB lower than the associated frequency component, wherein the compressed audio enhancement system comprises a downward expander having an attack time of less than one millisecond, a first system configured to receive an input audio data signal and to electronically process the input audio data signal to generate level normalized audio data, a second system configured to receive the level normalized audio data and to electronically process the level normalized audio data to generate a double

sideband AM signal, a third system configured to receive the double sideband AM signal and to mix the double sideband AM signal with the input audio data to generate primary audio data frequency components having double sideband AM signal side components, wherein the electronic processing comprises processing with a suitable combination of hardware and software, wherein the first system comprises a first automatic gain control processor configured to receive the input audio data signal to remove a DC component of the input audio data signal and a second automatic gain control processor configured to receive the DC-compensated audio data signal and the input audio data signal and to generate the level normalized audio data, wherein the second system comprises a Hilbert converter configured to receive the level normalized audio data and to electronically process the level normalized audio data to decorrelate the level normalized audio data and a downward compressor configured to receive the level normalized audio data and to generate the double sideband AM signal from the level normalized audio data, wherein the first system comprises a high pass filter coupled to the second automatic gain control processor and configured to receive the level normalized audio data and to generate high-pass filtered level normalized audio data, a low pass filter coupled to the high pass filter and configured to receive the high-pass filtered level normalized audio data and to generate band-pass filtered normalized audio data, a Hilbert transform configured to receive the band pass filtered normalized audio data and to form a Bedrosian all pass pair and a Butterworth filter coupled to the Hilbert transform and configured to remove low frequency components, a first automatic gain control unit having a first threshold value, a first target value, a first smoothing time and a first recovery rate, configured to receive an audio input and to generate an

audio output, a second automatic gain control unit having a second threshold value, a second target value, a second smoothing time and a second recovery rate, configured to receive the audio output of the first automatic gain control unit and to generate a second audio output, and wherein the first threshold value is less than the second threshold value, the first target value is less than the second target value, the first smoothing time is longer than the second smoothing time, the first recovery rate is essentially equal to the second recovery rate, a time of day unit configured to add one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a time of day, a programming type unit configured to add one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a programming type, a plurality of third automatic gain control units coupled between the first automatic gain control unit and the second automatic gain control unit, wherein each of the plurality of third automatic gain control units has a threshold value that is between the first threshold value and the second threshold value, wherein each of the plurality of third automatic gain control units has a target value that is between the first target value and the second target value, wherein each of the plurality of third automatic gain control units has a smoothing time that is between the first smoothing time and the second smoothing time, and wherein each of the plurality of third automatic gain control units has a recovery rate that is essentially equal to the first recovery rate and the second recovery rate, an audio data source, a delay system having a plurality of Z^{-N} delays coupled to the audio data source and a delay input and configured to delay audio data provided by the audio data source by an amount of time equal to the delay input,

a randomizer system coupled to the delay input and configured to generate an output that randomly changes as a function of time, a scattering matrix coupled to the audio data source and the delay system, an adder system coupled to the audio data source and to a feedback source derived from the delayed audio output, a low pass filter coupled to a delay output of the delay system, a scaler receiving the delayed audio data, scaling the audio data and providing the scaled and delayed audio data to a feedback path, a mixer coupled to the audio data source and configured to receive the delayed audio data and to mix the delayed audio data with non-delayed audio data from the audio data source, wherein the audio data source, the delay system and the randomizer system are digital data processing systems operating in the time domain, and wherein the randomizer system is configured to generate an output that randomly changes as a function of time by no more than 5 cents, and using the system to perform a method comprising receiving digitally encoded audio data at an audio processing system; modifying the digitally encoded audio data to add additional perceptually-masked audio data having an energy sufficient to prevent kinocilia of a listener from becoming dormant; generating sound waves with a sound wave generating device using the modified digitally encoded audio data; filtering low frequency components of the digitally encoded audio data from the digitally encoded audio data; filtering high frequency components of the digitally encoded audio data from the digitally encoded audio data; applying a Hilbert transform to the digitally encoded audio data; determining an absolute value of the digitally encoded audio data; adding modulation distortion to the digitally encoded audio data; processing the digitally encoded audio data with a downward expander having an attack time of less than 1 millisecond; adding modulation distortion to the digitally

encoded audio data having a magnitude of at least 13 dB less than a magnitude of an associated audio frequency component; receiving an input audio data signal at a first system and electronically processing the input audio data signal to generate level normalized audio data; receiving the level normalized audio data at a second system configured to and electronically processing the level normalized audio data to generate a double sideband AM signal; receiving the double sideband AM signal at a third system and mixing the double sideband AM signal with the input audio data to generate primary audio data frequency components having double sideband AM signal side components; wherein the electronic processing comprises processing with a suitable combination of hardware and software; removing a DC component of the input audio data signal using a first automatic gain control processor; receiving the DC-compensated audio data signal and the input audio data signal at a second automatic gain control processor and generating the level normalized audio data; receiving the level normalized audio data at a Hilbert converter and electronically processing the level normalized audio data to decorrelate the level normalized audio data; receiving the level normalized audio data at a downward compressor and generating the double sideband AM signal from the level normalized audio data; receiving the level normalized audio data at a high pass filter coupled to the second automatic gain control processor and generating high-pass filtered level normalized audio data; receiving the high-pass filtered level normalized audio data at a low pass filter coupled to the high pass filter and generating band-pass filtered normalized audio data; receiving the band pass filtered normalized audio data at a Hilbert transform and forming a Bedrosian all pass pair; receiving the audio input at the first automatic gain control unit; generating an audio output with the

first automatic gain control unit; receiving the audio output of the first automatic gain control unit at the second automatic gain control; generating the second audio output using the second automatic gain control unit; wherein the first threshold value is less than the second threshold value, the first target value is less than the second target value, the first smoothing time is longer than the second smoothing time, the first recovery rate is essentially equal to the second recovery rate; adding one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a time of day using a time of day unit; adding one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a programming type using a programming type unit; receiving encoded audio data at an audio processing system; generating an output that randomly changes as a function of time; delaying the encoded audio data by an amount of time equal to the output; processing the encoded audio data using a scattering matrix; adding the encoded audio data to a feedback source derived from the delayed encoded audio output; passing a low frequency component of the delayed encoded audio data; scaling the audio data; providing the scaled and delayed audio data to a feedback path; mixing the delayed encoded audio data with non-delayed encoded audio data from an audio data source; wherein generating the output that randomly changes as a function of time comprises applying a Z^{-N} delay; wherein generating the output that randomly changes as a function of time comprises applying a plurality of Z^{-N} delays; and wherein generating an output that randomly changes as a function of time comprises generating an output that randomly changes as a function of time by no more than 5 cents. While many other broader applications of the present disclosure are contemplated

as claimed below, at least this one narrow exemplary embodiment is also contemplated.

[00112] Audio data can include events that are louder than other events, such as gunshots, cymbal crashes, drum beats and so forth. When these events occur, they mask audio data that is 13 dB lower in gain for a period of time (typically around 200 milliseconds), such as audio data that has the same frequency components as the frequency components of the event. This masking occurs as a result of the psychoacoustic processes related to hearing. However, even though the masked audio signals cannot be perceived, the nerve cells in the organ of Corti are still receiving the masked audio signals, and are using energy to process them. This additional energy use results in a loss of hearing sensitivity. As such, the audio processing system that amplifies such signals is not only wasting energy on amplification of signals that are not perceived by the listener, it is also wasting that energy to create an inferior listening experience.

[00113] By detecting such transient events and dynamically equalizing the audio data to reduce the audio signals that will be masked, the amount of energy consumed by the audio processing system can be reduced, which can result in longer battery life. In addition, the effect of such masked audio signals on the nerves in the organ of Corti can be reduced or eliminated, which results in an improved audio experience for the listener.

[00114] **FIGURE 13** is a diagram of a system 1300 for dynamic equalization of audio data, in accordance with an exemplary embodiment of the present disclosure. System 1300 can be implemented in hardware or a suitable combination of hardware and software.

[00115] System 1300 includes crossover 1302, which receives audio data and processes the audio data to generate separate

frequency bands of audio data. In one exemplary embodiment, crossover 1302 can generate a first band having a frequency range of 0-50 Hz, a second band having a frequency range of 50-500 Hz, a third band having a frequency range of 500-4500 Hz and a fourth band having a frequency range of 4500 Hz and above, or other suitable numbers of bands and associated frequency ranges can also or alternatively be used. The input to crossover 1302 can be an unprocessed audio signal, a normalized audio signal or other suitable audio signals.

[00116] The outputs of crossover 1302 are further filtered using associated filters, such as low pass filter 1304, low mid pass filter 1306, mid pass filter 1308 and high pass filter 1310, or other suitable filters. In addition, the high frequency band can be further processed to add harmonic components, such as to compensate for lossy compression processing of the audio data that can result in audio data having a narrow image width and sparse frequency components. In one exemplary embodiment, the harmonic components can be added using clipping circuit 1312, which generates harmonic components by clipping the high frequency components of the audio data. High pass filter 1314 is used to remove lower frequency harmonic components, and scaler 1316 is used to control the magnitude of the harmonic processed audio that is added to the unprocessed audio at adder 1338. Control of scaler 1316 is provided by crossover 1318, which can generate a high frequency band output for frequencies above a predetermined level, such as 8000 Hz, and a low frequency band output for frequencies below the predetermined level. The RMS values of the high and low frequency bands are generated by RMS processors 1320 and 1322, and the RMS values are then converted from linear values to log values by DB20 1324 and DB20 1326, respectively. The difference between the high and low frequency components is then determined using subtractor 1328, and a value

from table 1330 is used to determine the amount of high frequency harmonic frequency component signal to be added to the unprocessed high frequency audio signal. In one exemplary embodiment, the amount can be set to zero until there is a 6 dB difference between the low and high frequency components, and as the difference increases from 6 dB to 10 dB, the amount of high frequency harmonic frequency component signal that is added to the unprocessed high frequency audio signal can increase from 0 dB to 8 dB. As the difference between the low and high frequency components increases from 10 dB to 15 dB, the amount of high frequency harmonic frequency component signal that is added to the unprocessed high frequency audio signal can increase from 8 dB to 9 dB. Likewise, other suitable amounts of high frequency harmonic frequency component signal can be added to the unprocessed high frequency audio signal. Increasing the amount of high frequency harmonic frequency component signal that is added to the unprocessed high frequency audio signal as a function of the change in the relative content of low and high frequency components of the high frequency band can be used to improve audio quality, because the difference is indicative of a sparse audio signal. The additional harmonic content helps to improve a sparse audio signal by providing additional frequency components that are complementary to the audio data. The high frequency harmonic components are then added to the unprocessed high frequency components by adder 1338.

[00117] Equalization of the low pass frequency component is accomplished using scaler 1332 under control of an input A, equalization of the mid pass frequency component is accomplished using scaler 1336 under control of an input B, and equalization of the high pass frequency component is accomplished using scaler 140 under control of an input C. The outputs of the

equalized audio components and the unprocessed low-mid pass output 1306 are combined using adder 1342 to generate an output.

[00118] In operation, system 1300 performs dynamic equalization to reduce power loss and also improves audio signal quality. The psychoacoustic masking processes result in a 150 to 200 millisecond loss in perception when a masking input having a crest of 13 dB or more is generated, due to the reaction of kinocilia to such audio inputs. When such transients occur, then maintain or increasing audio during the dead zone that follows the transient only serves to increase the power consumed by the audio processing system without increasing audio quality. In addition, while the audio input is not resulting in nerve signals that ultimately reach the listener's brain, processing of that audio energy still requires work to be done by kinocilia in the organ of Corti, and can also increase the amount of additional energy that is required in order to generate a perceptible response. By dynamically equalizing the audio data to reduce the gain during such periods, system 1300 helps to reduce both the amount of energy required to process the audio data as well as the amount of energy required by the listener to listen to the audio data. In addition, adding harmonic frequency content to the high frequency audio data when the total audio data content is sparse helps to improve the perceived audio quality, by providing additional frequency components in the sparse audio data that complement the existing frequency components.

[00119] **FIGURE 14** is a diagram of a system 1400 for controlling dynamic equalization of audio data, in accordance with an exemplary embodiment of the present disclosure. System 1400 can be implemented in hardware or a suitable combination of hardware and software.

[00120] System 1400 includes automatic gain control core 1402 and automatic gain control multiplier 1404, which are configured to receive an audio signal input and to generate a normalized audio signal output. The normalized audio signal output of AGC multiplier 1404 can also be provided to crossover 1302 or other suitable systems or components.

[00121] Filter 1406 can be a band pass filter having a frequency range of 40 to 80 Hz or other suitable filters. The output from filter 1406 is processed by RMS processor 1408 to generate a signal that represents the RMS value of the output of filter 1406. Derivative processor 1410 receives the band pass RMS value and generates an output that represents the rate of change of the band pass RMS value. Downward expander 1412 is used to prevent dynamic equalization of the associated frequency band when there is no associated transient occurring in the frequency band.

[00122] Filter 1414 can be a band pass filter having a frequency range of 500 to 4000 Hz or other suitable filters. The output from filter 1414 is processed by RMS processor 1416 to generate a signal that represents the RMS value of the output of filter 1414. Derivative processor 1418 receives the band pass RMS value and generates an output that represents the rate of change of the band pass RMS value. Downward expander 1420 is used to prevent dynamic equalization of the associated frequency band when there is no associated transient occurring in the frequency band.

[00123] Filter 1422 can be a high pass filter having a frequency range of 4000 Hz and above or other suitable filters. The output from filter 1422 is processed by RMS processor 1424 to generate a signal that represents the RMS value of the output of filter 1422. Derivative processor 1426 receives the band pass RMS value and generates an output that represents the rate of

change of the band pass RMS value. Downward expander 1428 is used to prevent dynamic equalization of the associated frequency band when there is no associated transient occurring in the frequency band.

[00124] In operation, system 1400 generates control inputs for a dynamic equalizer, by detecting transients in frequency bands associated with the dynamic equalization. System 1400 thus helps to improve power consumption for an audio data processor, and also helps to improve the perceptual audio quality.

[00125] **FIGURE 15** is a diagram of an algorithm 1500 for dynamic equalization of audio data, in accordance with an exemplary embodiment of the present disclosure. Algorithm 1500 can be implemented in hardware or a suitable combination of hardware and software. Although algorithm 1500 is shown as a flow chart, other suitable programming paradigms such as state diagrams and object-oriented programming can also or alternatively be used.

[00126] Algorithm 1500 begins at 1502, where audio data is received and processed, such as to generate a normalized audio signal by using a first adaptive gain control processor that is used to remove a DC signal component and a second adaptive gain control processor that receives the output of the first adaptive gain control processor and the input audio, or in other suitable manners. The algorithm then proceeds to 1504, where the audio data is filtered to generate different bands of audio data. The algorithm then proceeds in parallel to 1506 and 1514.

[00127] At 1506, the high and low frequency components of one or more of the bands of audio data are analyzed, such as to determine whether there is a greater RMS value of one component compared to the other. The algorithm then proceeds to 1508, where it is determined whether the difference is indicative of an audio signal that is sparse or that otherwise would benefit

from additional harmonic content, such as by exceeding a predetermined level. If the difference does not exceed the level, the algorithm proceeds to 1518 where the unprocessed signal is dynamically equalized, otherwise the algorithm proceeds to 1510, where harmonic content is generated. In one exemplary embodiment, harmonic content can be generated by clipping the audio signal and then filtering the clipped signal to remove predetermined harmonic frequency components, or in other suitable manners. The algorithm then proceeds to 1512, where the harmonic content is added to unprocessed audio signal and the combined signal is processed at 1518.

[00128] At 1514, the audio signals are processed to determine whether a transient has occurred, such as by generating a derivative of an RMS value of the frequency component or in other suitable manners. A downward expander or other suitable components or processes can be used to ensure that the control signal is only generated for significant transients that will cause an associated psychoacoustic masking of predetermined associated audio frequency components. The algorithm then proceeds to 1516, where a control signal is generated based on the transient. The control signal is applied to the audio signal to perform dynamic equalization at 1518, such as to reduce the gain of the audio signal frequency components when a masking transient occurs, to both reduce power consumption and listener fatigue, and to improve audio quality to the listener.

[00129] In operation, algorithm 1500 allows audio data to be dynamically equalized, by detecting masking transients and by using such masking transients to generate dynamic equalization controls. By dynamically equalizing the audio data that would otherwise not be perceptible to the listener, the amount of energy required to process the audio data can be reduced, and the perceived quality of the audio data can be improved.

[00130] **FIGURE 16** is a diagram of a system 1600 for parametric stereo processing, in accordance with an exemplary embodiment of the present disclosure. System 1600 can be implemented in hardware or a suitable combination of hardware and software.

[00131] System 1600 includes time to frequency conversion system 1602, which converts frames of a time-varying audio signal into frames of frequency components, such as by performing a fast-Fourier transform or in other suitable manners.

[00132] Bin comparison system 1604 receives the frames of frequency domain data and compares the magnitude of the left channel audio data with the magnitude of the right channel audio data for each frequency bin.

[00133] Phase adjustment system 1606 receives the comparison data for each bin of frequency data from bin comparison system 1604 and sets the phase of the right channel frequency bin component equal to the phase of the left channel frequency bin component if the magnitude of the left channel frequency bin component is greater than the magnitude of the right channel frequency bin component. The output of phase adjustment system 406 is parametric audio data.

[00134] Surround processing system 1608 receives the parametric audio data and generates surround audio data. In one exemplary embodiment, surround processing system 1608 can receive speaker location data and can calculate a phase angle difference for the input audio data that corresponds to the location of the speaker. In this exemplary embodiment, surround processing system 1608 can generate audio data for any suitable number of speakers in any suitable locations, by adjusting the phase angle of the parametric audio data to reflect the speaker location relative to other speakers.

[00135] In operation, system 1600 allows audio data to be processed to generate parametric audio data, which can then be processed based on predetermined speaker locations to generate N-dimensional audio data. System 1600 eliminates the phase data of the input audio data, which is not needed when the input audio data is processed to be output from speakers in non-stereophonic speaker locations.

[00136] FIGURE 17 is a diagram of an algorithm 1700 for parametric stereo processing, in accordance with an exemplary embodiment of the present disclosure. Algorithm 1700 can be implemented in hardware or a suitable combination of hardware and software. Although algorithm 1700 is shown as a flow chart, other suitable programming paradigms such as state diagrams and object-oriented programming can also or alternatively be used.

[00137] Algorithm 1700 begins at 1702 where audio data is received, such as analog or digital audio data in the time domain. The algorithm then proceeds to 1704.

[00138] At 1704, the audio data is converted from the time domain to the frequency domain, such as by performing a fast Fourier transform on the audio data or in other suitable manners. The algorithm then proceeds to 1706.

[00139] At 1706, it is determined whether the magnitude of a left channel frequency component of the frequency domain audio data is greater than the magnitude of the associated right channel frequency component. In one exemplary embodiment, 1706 can be performed on a frequency component basis for each of the frequency components of the audio data, or in other suitable manners. If it is determined that the magnitude of a left channel frequency component of the frequency domain audio data is not greater than the magnitude of the associated right channel frequency component, the algorithm proceeds to 1710, otherwise the algorithm proceeds to 1708, where the phase of the

right channel frequency component is replaced with the phase of the left channel frequency component. The algorithm then proceeds to 1710.

[00140] At 1710, the audio data is processed for an N-channel surround playback environment. In one exemplary embodiment, the locations of each of a plurality of speakers can be input into a system which can then determine a preferred phase relationship of the left and right channel audio data for that speaker. The phase and magnitude of the audio data can then be generated as a function of the speaker location, or other suitable processes can also or alternatively be used.

[00141] **FIGURE 18** is a diagram of a system 1800 for providing artifact masking of audio data artifacts in accordance with an exemplary embodiment of the present disclosure. System 1800 can be implemented in hardware or a suitable combination of hardware and software, and can be one or more software systems operating on a special purpose audio processor or other suitable devices.

[00142] System 1800 includes AGC core 1802, which provides automatic gain control of input audio data, which can be digital audio data in the time domain or other suitable audio data. AGC core 1802 provides averaged input audio data over a 200 millisecond period or other suitable periods. The output of AGC core 1802 is provided to AGC multiplier 1804 with the input audio data, and which normalizes the input audio data. The output of AGC multiplier 1804 is then provided to de-interleaver 1806, which can separate the audio data into left and right channels or other suitable numbers of channels, if multiple channels are present. Subtractor 1808 subtracts the left channel from the right channel and maximum absolute value detector 1810 generates an output for downward expander 1812. The output of maximum absolute value detector 1810 provides an indication of the level of audio artifacts in the audio data, as highly

compressed audio with a large number of audio artifacts will also have a large value of L-R.

[00143] Downward expander 1812 can have a very short attack time such as 0.01 milliseconds or less, and a decay time that ranges from 10 to 80 milliseconds, such as 30 milliseconds. This signal is provided to AGC multiplier 1814 with the low frequency signal components from reverb 1820 (discussed below). The output of AGC multiplier 1814 is the relative magnitude of the envelope of the L-R to L+R ratio. When a "wide" signal is being processed, which is one where the value of L-R is greater than the value of L+R, then downward expander 1812 has a gain of zero, so that when the ratio of L-R to L+R is 1 (which is representative of a normal stereo signal), AGC multiplier 1814 is provided with a control signal that turns it all the way on. This configuration results in a maximum injection of scaled, decorrelated low pass stereo. Thus, where the L-R component is large, which usually results in the generation of more audio artifacts, the maximum amount of auxiliary masker signals will be injected into the L-R signal.

[00144] More audio artifacts are typically generated with a large L-R value, such as when the number of bits of audio data that still need to be encoded as the encoder approaches the end of a frame, the codec looks for ways to save the number of bits that need to be encoded. In this case, some frequency lines might get set to zero (deleted), typically high frequency content. Furthermore, when the loss of bandwidth changes from frame to frame, the generation of artifacts can become more noticeable. For example, encoders can switch between mid/side stereo coding and intensity stereo coding for lower bit-rates, but switching between modes can result in the generation of audio artifacts.

[00145] The output of AGC multiplier 1814 is provided to de-interleaver 1826, which separates the signal into left and right channel components. The left channel component is provided to Hilbert transform 1830, and the right channel component is inverted by inverter 1832 and is provided to Hilbert transform 1834.

[00146] The input audio data is also provided to crossover filter 1816 which generates a low frequency component that is provided to scaler 1818 and adder 1822, and a high frequency component that is also added to adder 1822. The use of crossover 1816 and adder 1822 generates audio data having the same phase as the low pass filter by itself, and helps to synchronize the processed L-R audio and the unprocessed L-R audio. The sum has a flat magnitude response, where the phase is at 90 or 270 shift relative to crossover 1816. Low pass filter 1824 removes high frequency components and the resulting signal is provided to de-interleaver 1828, which generates a left channel output to Hilbert transform 1830 and a right channel output to Hilbert transform 1834.

[00147] The low frequency output from crossover filter 1816 is provided to scaler 1818, which scales the low frequency output and provides the low frequency output to reverb 1820, which can utilize a variable time delay unit such as that disclosed herein. The decorrelated low frequency audio data mixes with the output of downward expander 1812, as discussed. The left channel components provided to Hilbert filter 1830 are added by adder 1836, and the right channel frequency components provided to Hilbert filter 1834 are added by adder 1838. The outputs of adders 1836 and 1838 are then provided to interleaver 1840 to generate an output audio signal. Hilbert filter 1830 receives the low pass, decorrelated multiplied version of L-R at a 90 degree-shifted input, and unprocessed L-R audio at 0 degrees,

which sums the negative components and will not cause the audio image to shift to the left. Hilbert filter 1834 generates synthetic artifact masking by moving the artifacts to the back of the image, where other artifacts are, and not in the front stereo image, where the audio artifacts will be more readily apparent.

[00148] In operation, system 1800 eliminates audio artifacts that are generated when audio data that has a large L-R to L+R ratio is being processed, by generating masking signals that cause the audio artifacts to be relocated out of the front stereo audio and moved into the surround stereo audio.

[00149] **FIGURE 19** is a diagram of an algorithm 1900 for processing audio data in accordance with an exemplary embodiment of the present disclosure. Algorithm 1900 can be implemented in hardware or a suitable combination of hardware and software, and can be one or more software systems operating on a special purpose audio processor.

[00150] Algorithm 1900 begins at 1902 and 1934 in parallel. At 1902, input audio data is separated into low and high frequency components, and then proceeds in parallel to 1904 and 1910. At 1904, the low and high frequency components are added back together, and at 1906, the combined low and high frequency components are processed with a low pass filter. The left and right channel components of the audio data are then separated.

[00151] At 1910, the low frequency components are scaled and are then decorrelated at 1912, such as by processing through a reverb processor or in other suitable manners.

[00152] At 1924, the input audio is processed to remove DC components of the audio data and to average the audio data over 200 millisecond periods. The algorithm then proceeds to 1926 where the level is normalized, and then the left and right channels are separated at 1928, such as with a de-interleaver or

in other suitable manners. At 1930, the difference between the left and right channel audio data is determined, such as by subtraction or in other suitable manners, and the algorithm proceeds to 1932 where the maximum absolute value of the difference signal is determined. The algorithm then proceeds to 1934 where a mask control signal is generated, such as by using a downward expander or in other suitable manners. The mask control signal is then multiplied at 1914 with the decorrelated low frequency audio signal components, and is separated into left and right channel signals at 1916.

[00153] At 1918, the left channel components from 1908 and 1916 are added, either directly or after other suitable processing, such as inversion, and at 1920, the right channel components from 1908 and 1916 are added, either directly or after other suitable processing, such as inversion. The combined left and right audio channel inputs are then used to generate an audio output signal having improved masking of audio artifacts.

[00154] A system for processing digitally-encoded audio data is disclosed that comprises a compressed audio source device providing a sequence of frames of compressed digital audio data, a compressed audio enhancement system configured to receive the sequence of frames of compressed digital audio data and to generate enhanced audio data by adding masked digital audio data to the sequence of frames of compressed digital audio data, where the masked digital audio data has an energy level sufficient to keep a kinocilia of a listener active and one or more speakers configured to receive the enhanced audio data and to generate sound waves using the enhanced audio data.

[00155] The system can further comprise a high pass filter configured to remove low frequency components of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data.

[00156] The system can further comprise a low pass filter configured to remove high frequency components of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data.

[00157] The system can further comprise a Hilbert transform configured to apply a phase shift to the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data.

[00158] The system can further comprise an absolute value processor configured to generate an absolute value of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data.

[00159] The system can further comprise generating modulation distortion of the enhanced audio data.

[00160] The system generating modulation distortion for one or more frequency components of the enhanced audio data, wherein the modulation distortion has a magnitude at least 13 dB lower than the associated frequency component.

[00161] The system can further comprise generating modulation distortion for one or more frequency components of the enhanced audio data, the modulation distortion having a frequency range centered at each of the associated frequency components, wherein the modulation distortion has a magnitude at least 13 dB lower than the associated frequency component.

[00162] The system can further comprise a downward expander.

[00163] The system can further comprise a downward expander having an attack time of less than one millisecond.

[00164] A method for processing digitally-encoded audio data is disclosed that comprises receiving digitally encoded audio data at an audio processing system, modifying the digitally encoded audio data to add additional perceptually-masked audio data having an energy sufficient to prevent kinocilia of a

listener from becoming dormant and generating sound waves with a sound wave generating device using the modified digitally encoded audio data.

[00165] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises filtering low frequency components of the digitally encoded audio data from the digitally encoded audio data.

[00166] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises filtering high frequency components of the digitally encoded audio data from the digitally encoded audio data.

[00167] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises applying a Hilbert transform to the digitally encoded audio data.

[00168] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises determining an absolute value of the digitally encoded audio data.

[00169] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises adding modulation distortion to the digitally encoded audio data.

[00170] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises processing the digitally encoded audio data with a downward expander.

[00171] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises processing the digitally encoded audio data with a downward expander having an attack time of less than 1 millisecond.

[00172] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises adding modulation distortion to the digitally encoded audio data having a magnitude of at least 13 dB less than a magnitude of an associated audio frequency component.

[00173] The method can further comprise wherein modifying the digitally encoded audio data to add the additional perceptually-masked audio data having the energy sufficient to prevent the kinocilia of the listener from becoming dormant comprises adding modulation distortion to the digitally encoded audio data having a magnitude less than a magnitude of an associated audio frequency component.

[00174] A system for processing audio data is disclosed that includes a first system configured to receive an input audio data signal and to electronically process the input audio data signal to generate level normalized audio data, a second system configured to receive the level normalized audio data and to electronically process the level normalized audio data to

generate a double sideband AM signal and a third system configured to receive the double sideband AM signal and to mix the double sideband AM signal with the input audio data to generate primary audio data frequency components having double sideband AM signal side components.

[00175] The system can further comprise processing with a suitable combination of hardware and software.

[00176] The system can further comprise a first automatic gain control processor configured to receive the input audio data signal to remove a DC component of the input audio data signal.

[00177] The system can further comprise a second automatic gain control processor configured to receive the DC-compensated audio data signal and the input audio data signal and to generate the level normalized audio data.

[00178] The system can further comprise a Hilbert converter configured to receive the level normalized audio data and to electronically process the level normalized audio data to decorrelate the level normalized audio data.

[00179] The system can further comprise a downward compressor configured to receive the level normalized audio data and to generate the double sideband AM signal from the level normalized audio data.

[00180] The system can further comprise a high pass filter coupled to the second automatic gain control processor and configured to receive the level normalized audio data and to generate high-pass filtered level normalized audio data.

[00181] The system can further comprise a low pass filter coupled to the high pass filter and configured to receive the high-pass filtered level normalized audio data and to generate band-pass filtered normalized audio data.

[00182] The system can further comprise a Hilbert transform configured to receive the band pass filtered normalized audio data and to form a Bedrosian all pass pair.

[00183] The system can further comprise a Butterworth filter coupled to the Hilbert transform and configured to remove low frequency components.

[00184] A method for processing audio data is disclosed that comprises receiving an input audio data signal at a first system and electronically processing the input audio data signal to generate level normalized audio data, receiving the level normalized audio data at a second system configured to and electronically processing the level normalized audio data to generate a double sideband AM signal and receiving the double sideband AM signal at a third system and mixing the double sideband AM signal with the input audio data to generate primary audio data frequency components having double sideband AM signal side components.

[00185] The method can further comprise processing with a suitable combination of hardware and software.

[00186] The method can further comprise removing a DC component of the input audio data signal using a first automatic gain control processor.

[00187] The method can further comprise receiving the DC-compensated audio data signal and the input audio data signal at a second automatic gain control processor and generating the level normalized audio data.

[00188] The method can further comprise receiving the level normalized audio data at a Hilbert converter and electronically processing the level normalized audio data to decorrelate the level normalized audio data.

[00189] The method can further comprise receiving the level normalized audio data at a downward compressor and generating

the double sideband AM signal from the level normalized audio data.

[00190] The method can further comprise receiving the level normalized audio data at a high pass filter coupled to the second automatic gain control processor and generating high-pass filtered level normalized audio data.

[00191] The method can further comprise receiving the high-pass filtered level normalized audio data at a low pass filter coupled to the high pass filter and generating band-pass filtered normalized audio data.

[00192] The method can further comprise receiving the band pass filtered normalized audio data at a Hilbert transform and forming a Bedrosian all pass pair.

[00193] The method can further comprise receiving the band pass filtered normalized audio data and forming a Bedrosian all pass pair.

[00194] A system for controlling a volume level of television program audio data is disclosed that comprises a first automatic gain control unit having a first threshold value, a first target value, a first smoothing time and a first recovery rate, configured to receive an audio input and to generate an audio output, a second automatic gain control unit having a second threshold value, a second target value, a second smoothing time and a second recovery rate, configured to receive the audio output of the first automatic gain control unit and to generate a second audio output and wherein the first threshold value is less than the second threshold value.

[00195] The system can further comprise wherein the first target value is less than the second target value.

[00196] The system can further comprise wherein the first smoothing time is longer than the second smoothing time.

[00197] The system can further comprise wherein the first recovery rate is essentially equal to the second recovery rate.

[00198] The system can further comprise a time of day unit configured to add one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a time of day.

[00199] The system can further comprise a programming type unit configured to add one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a programming type.

[00200] The system can further comprise a plurality of third automatic gain control units coupled between the first automatic gain control unit and the second automatic gain control unit.

[00201] The system can further comprise wherein each of the plurality of third automatic gain control units has a threshold value that is between the first threshold value and the second threshold value.

[00202] The system can further comprise wherein each of the plurality of third automatic gain control units has a target value that is between the first target value and the second target value.

[00203] The system can further comprise wherein each of the plurality of third automatic gain control units has a smoothing time that is between the first smoothing time and the second smoothing time.

[00204] The system can further comprise wherein each of the plurality of third automatic gain control units has a recovery rate that is essentially equal to the first recovery rate and the second recovery rate.

[00205] A method for controlling a volume level of television program audio data is disclosed that comprises receiving an audio input at a first automatic gain control unit having a first threshold value, a first target value, a first smoothing time and a first recovery rate, generating an audio output with the first automatic gain control unit using the first threshold value, the first target value, the first smoothing time and the first recovery rate, receiving the audio output of the first automatic gain control unit at a second automatic gain control unit having a second threshold value, a second target value, a second smoothing time and a second recovery rate, generating a second audio output using the second automatic gain control unit having the second threshold value, the second target value, the second smoothing time and the second recovery rate and wherein the first threshold value is less than the second threshold value.

[00206] The method can further comprise wherein the first target value is less than the second target value.

[00207] The method can further comprise wherein the first smoothing time is longer than the second smoothing time.

[00208] The method can further comprise wherein the first recovery rate is essentially equal to the second recovery rate.

[00209] The method can further comprise adding one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a time of day using a time of day unit.

[00210] The method can further comprise adding one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a programming type using a programming type unit.

[00211] The method can further comprise a plurality of third automatic gain control units coupled between the first automatic gain control unit and the second automatic gain control unit.

[00212] The method can further comprise wherein each of the plurality of third automatic gain control units has a threshold value that is between the first threshold value and the second threshold value.

[00213] The method can further comprise wherein each of the plurality of third automatic gain control units has a predetermined threshold value.

[00214] A system for processing audio data is disclosed that comprises an audio data source, a delay system coupled to the audio data source and a delay input and configured to delay audio data provided by the audio data source by an amount of time equal to the delay input and a randomizer system coupled to the delay input and configured to generate an output that randomly changes as a function of time.

[00215] The system can further comprise wherein the delay system comprises a Z^{-N} delay.

[00216] The system can further comprise wherein the delay system comprises a plurality of Z^{-N} delays.

[00217] The system can further comprise a scattering matrix coupled to the audio data source and the delay system.

[00218] The system can further comprise an adder system coupled to the audio data source and to a feedback source derived from the delayed audio output.

[00219] The system can further comprise a low pass filter coupled to a delay output of the delay system.

[00220] The system can further comprise a scaler receiving the delayed audio data, scaling the audio data and providing the scaled and delayed audio data to a feedback path.

[00221] The system can further comprise a mixer coupled to the audio data source and configured to receive the delayed audio data and to mix the delayed audio data with non-delayed audio data from the audio data source.

[00222] The system can further comprise wherein the audio data source, the delay system and the randomizer system are digital data processing systems operating in the time domain.

[00223] The system can further comprise wherein the randomizer system is configured to generate an output that randomly changes as a function of time by no more than 5 cents.

[00224] A method for processing audio data is disclosed that comprises receiving encoded audio data at an audio processing system, generating an output that randomly changes as a function of time and delaying the encoded audio data by an amount of time equal to the output.

[00225] The method can further comprise wherein generating the output that randomly changes as a function of time comprises applying a Z^{-N} delay.

[00226] The method can further comprise wherein generating the output that randomly changes as a function of time comprises applying a plurality of Z^{-N} delays.

[00227] The method can further comprise processing the encoded audio data using a scattering matrix.

[00228] The method can further comprise adding the encoded audio data to a feedback source derived from the delayed encoded audio output.

[00229] The method can further comprise passing a low frequency component of the delayed encoded audio data.

[00230] The method can further comprise scaling the audio data and providing the scaled and delayed audio data to a feedback path.

[00231] The method can further comprise mixing the delayed encoded audio data with non-delayed encoded audio data from an audio data source.

[00232] The method can further comprise generating an output that randomly changes as a function of time comprises generating an output that randomly changes as a function of time by no more than 5 cents.

[00233] In a system for processing digitally-encoded audio data that has a compressed audio source device providing a sequence of frames of compressed digital audio data, a compressed audio enhancement system configured to receive the sequence of frames of compressed digital audio data and to generate enhanced audio data by adding masked digital audio data to the sequence of frames of compressed digital audio data, where the masked digital audio data has an energy level sufficient to keep a kinocilia of a listener active, one or more speakers configured to receive the enhanced audio data and to generate sound waves using the enhanced audio data, a high pass filter configured to remove low frequency components of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, a low pass filter configured to remove high frequency components of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, a Hilbert transform configured to apply a phase shift to the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, an absolute value processor configured to generate an absolute value of the sequence of frames of compressed digital audio data prior to generation of the enhanced audio data, wherein generation of the enhanced audio data comprises generating modulation distortion of the enhanced audio data, wherein generation of the enhanced audio data comprises generating modulation distortion for one or more frequency components of the enhanced audio data, wherein generation of the enhanced audio data comprises generating modulation distortion for one or more frequency components of the enhanced audio data, the modulation distortion having a frequency range centered at each of the associated frequency components, wherein the modulation distortion has a magnitude at least 13 dB lower than

the associated frequency component, wherein the compressed audio enhancement system comprises a downward expander having an attack time of less than one millisecond, a first system configured to receive an input audio data signal and to electronically process the input audio data signal to generate level normalized audio data, a second system configured to receive the level normalized audio data and to electronically process the level normalized audio data to generate a double sideband AM signal, a third system configured to receive the double sideband AM signal and to mix the double sideband AM signal with the input audio data to generate primary audio data frequency components having double sideband AM signal side components, wherein the electronic processing comprises processing with a suitable combination of hardware and software, wherein the first system comprises a first automatic gain control processor configured to receive the input audio data signal to remove a DC component of the input audio data signal and a second automatic gain control processor configured to receive the DC-compensated audio data signal and the input audio data signal and to generate the level normalized audio data, wherein the second system comprises a Hilbert converter configured to receive the level normalized audio data and to electronically process the level normalized audio data to decorrelate the level normalized audio data and a downward compressor configured to receive the level normalized audio data and to generate the double sideband AM signal from the level normalized audio data, wherein the first system comprises a high pass filter coupled to the second automatic gain control processor and configured to receive the level normalized audio data and to generate high-pass filtered level normalized audio data, a low pass filter coupled to the high pass filter and configured to receive the high-pass filtered level normalized

audio data and to generate band-pass filtered normalized audio data, a Hilbert transform configured to receive the band pass filtered normalized audio data and to form a Bedrosian all pass pair and a Butterworth filter coupled to the Hilbert transform and configured to remove low frequency components, a first automatic gain control unit having a first threshold value, a first target value, a first smoothing time and a first recovery rate, configured to receive an audio input and to generate an audio output, a second automatic gain control unit having a second threshold value, a second target value, a second smoothing time and a second recovery rate, configured to receive the audio output of the first automatic gain control unit and to generate a second audio output, and wherein the first threshold value is less than the second threshold value, the first target value is less than the second target value, the first smoothing time is longer than the second smoothing time, the first recovery rate is essentially equal to the second recovery rate, a time of day unit configured to add one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a time of day, a programming type unit configured to add one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a programming type, a plurality of third automatic gain control units coupled between the first automatic gain control unit and the second automatic gain control unit, wherein each of the plurality of third automatic gain control units has a threshold value that is between the first threshold value and the second threshold value, wherein each of the plurality of third automatic gain control units has a target value that is between the first target value and the second target value, wherein each of the plurality of third automatic gain control units has a

smoothing time that is between the first smoothing time and the second smoothing time, and wherein each of the plurality of third automatic gain control units has a recovery rate that is essentially equal to the first recovery rate and the second recovery rate, an audio data source, a delay system having a plurality of Z^{-N} delays coupled to the audio data source and a delay input and configured to delay audio data provided by the audio data source by an amount of time equal to the delay input, a randomizer system coupled to the delay input and configured to generate an output that randomly changes as a function of time, a scattering matrix coupled to the audio data source and the delay system, an adder system coupled to the audio data source and to a feedback source derived from the delayed audio output, a low pass filter coupled to a delay output of the delay system, a scaler receiving the delayed audio data, scaling the audio data and providing the scaled and delayed audio data to a feedback path, a mixer coupled to the audio data source and configured to receive the delayed audio data and to mix the delayed audio data with non-delayed audio data from the audio data source, wherein the audio data source, the delay system and the randomizer system are digital data processing systems operating in the time domain, and wherein the randomizer system is configured to generate an output that randomly changes as a function of time by no more than 5 cents, a method is disclosed comprising receiving digitally encoded audio data at an audio processing system, modifying the digitally encoded audio data to add additional perceptually-masked audio data having an energy sufficient to prevent kinocilia of a listener from becoming dormant, generating sound waves with a sound wave generating device using the modified digitally encoded audio data, filtering low frequency components of the digitally encoded audio data from the digitally encoded audio data, filtering high

frequency components of the digitally encoded audio data from the digitally encoded audio data, applying a Hilbert transform to the digitally encoded audio data, determining an absolute value of the digitally encoded audio data, adding modulation distortion to the digitally encoded audio data, processing the digitally encoded audio data with a downward expander having an attack time of less than 1 millisecond, adding modulation distortion to the digitally encoded audio data having a magnitude of at least 13 dB less than a magnitude of an associated audio frequency component, receiving an input audio data signal at a first system and electronically processing the input audio data signal to generate level normalized audio data, receiving the level normalized audio data at a second system configured to and electronically processing the level normalized audio data to generate a double sideband AM signal, receiving the double sideband AM signal at a third system and mixing the double sideband AM signal with the input audio data to generate primary audio data frequency components having double sideband AM signal side components, wherein the electronic processing comprises processing with a suitable combination of hardware and software, removing a DC component of the input audio data signal using a first automatic gain control processor, receiving the DC-compensated audio data signal and the input audio data signal at a second automatic gain control processor and generating the level normalized audio data, receiving the level normalized audio data at a Hilbert converter and electronically processing the level normalized audio data to decorrelate the level normalized audio data, receiving the level normalized audio data at a downward compressor and generating the double sideband AM signal from the level normalized audio data, receiving the level normalized audio data at a high pass filter coupled to the second automatic gain control processor and generating high-pass

filtered level normalized audio data, receiving the high-pass filtered level normalized audio data at a low pass filter coupled to the high pass filter and generating band-pass filtered normalized audio data, receiving the band pass filtered normalized audio data at a Hilbert transform and forming a Bedrosian all pass pair, receiving the audio input at the first automatic gain control unit, generating an audio output with the first automatic gain control unit, receiving the audio output of the first automatic gain control unit at the second automatic gain control, generating the second audio output using the second automatic gain control unit, wherein the first threshold value is less than the second threshold value, the first target value is less than the second target value, the first smoothing time is longer than the second smoothing time, the first recovery rate is essentially equal to the second recovery rate, adding one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a time of day using a time of day unit, adding one or more automatic gain control units to the first automatic gain control unit and the second automatic gain control unit as a function of a programming type using a programming type unit, receiving encoded audio data at an audio processing system, generating an output that randomly changes as a function of time, delaying the encoded audio data by an amount of time equal to the output, processing the encoded audio data using a scattering matrix, adding the encoded audio data to a feedback source derived from the delayed encoded audio output, passing a low frequency component of the delayed encoded audio data, scaling the audio data, providing the scaled and delayed audio data to a feedback path, mixing the delayed encoded audio data with non-delayed encoded audio data from an audio data source, wherein generating the output that randomly changes as a

function of time comprises applying a Z^{-N} delay, wherein generating the output that randomly changes as a function of time comprises applying a plurality of Z^{-N} delays and wherein generating an output that randomly changes as a function of time comprises generating an output that randomly changes as a function of time by no more than 5 cents.

[00234] It should be emphasized that the above-described embodiments are merely examples of possible implementations. Many variations and modifications may be made to the above-described embodiments without departing from the principles of the present disclosure. All such modifications and variations are intended to be included herein within the scope of this disclosure and protected by the following claims.

WHAT IS CLAIMED IS:

1. A system for processing audio data comprising:
 - a plurality of gain adjustment devices, each gain adjustment device having an associated audio input frequency band;
 - a plurality of control signal processing systems, each control signal processing system configured to receive audio input data for one of the associated audio input frequency bands and to generate a gain adjustment device control signal;wherein the gain adjustment device control signal is configured to decrease a gain setting of an associated gain adjustment device for a predetermined period of time as a function of a transient in the associated audio input frequency band.
2. The system of claim 1 further comprising a clipping system configured to receive the audio input data and to generate a clipped audio input signal.
3. The system of claim 2 further comprising a high pass filter coupled to the clipping system and configured to receive the clipped audio input signal and to generate high-pass filtered clipped audio data.
4. The system of claim 3 further comprising a scaler multiplier coupled to the high pass filter and configured to multiply the high-pass filtered clipped audio data by a predetermined value.
5. The system of claim 4 further comprising a table coupled to the scaler multiplier and configured to receive a difference signal and to select the predetermined value as a function of the difference signal.
6. The system of claim 5 further comprising a crossover filter coupled to the subtractor, the crossover filter

configured to generate a low frequency output and a high frequency output and the subtractor configured to receive the low frequency output and the high frequency output and to generate the difference signal.

7. The system of claim 6 further comprising an RMS converter coupled between the crossover filter and the subtractor.

8. The system of claim 6 further comprising a linear to log converter coupled between the crossover filter and the subtractor.

9. The system of claim 4 further comprising an adder coupled to the table and one of the gain adjustment devices.

10. A method for processing audio data comprising:
receiving audio input data for an associated audio input frequency band at a control signal processing system;
generating a gain adjustment device control for a gain adjustment device;
adjusting the gain adjustment device using the gain adjustment device control; and
decreasing the gain setting of the gain adjustment device for a predetermined period of time as a function of a transient in the associated audio input frequency band.

11. The method of claim 10 further comprising generating a clipped audio input signal from the audio input data with a clipping system.

12. The method of claim 11 further comprising generating high-pass filtered clipped audio data from the clipped audio input signal using a high pass filter coupled to the clipping system.

13. The method of claim 12 further comprising multiplying the high-pass filtered clipped audio data by a predetermined value using a scaler multiplier coupled to the high pass filter.

14. The method of claim 13 further comprising receiving a difference signal at a table coupled to the scaler multiplier and selecting the predetermined value as a function of the difference signal.

15. The method of claim 14 further comprising:

generating a low frequency output and a high frequency output with a crossover filter coupled to the subtractor; and

receiving the low frequency output and the high frequency output at the subtractor and generating the difference signal.

16. The method of claim 15 further comprising coupling an RMS converter between the crossover filter and the subtractor.

17. The method of claim 15 further comprising coupling a linear to log converter between the crossover filter and the subtractor.

18. The method of claim 13 further comprising coupling an adder to the table and one of the gain adjustment devices.

19. In a system for processing audio data having a plurality of gain adjustment devices, each gain adjustment device having an associated audio input frequency band, a plurality of control signal processing systems, each control signal processing system configured to receive audio input data for one of the associated audio input frequency bands and to generate a gain adjustment device control signal, wherein the gain adjustment device control signal is configured to decrease a gain setting of an associated gain adjustment device for a predetermined period of time as a function of a transient in the associated audio input frequency band, a clipping system configured to receive the audio input data and to generate a clipped audio input signal, a high pass filter coupled to the clipping system and configured to receive the clipped audio input signal and to generate high-pass filtered clipped audio data, a scaler multiplier coupled to the high pass filter and

configured to multiply the high-pass filtered clipped audio data by a predetermined value, a table coupled to the scaler multiplier and configured to receive a difference signal and to select the predetermined value as a function of the difference signal, a crossover filter coupled to the subtractor, the crossover filter configured to generate a low frequency output and a high frequency output and the subtractor configured to receive the low frequency output and the high frequency output and to generate the difference signal, an RMS converter coupled between the crossover filter and the subtractor, a linear to log converter coupled between the crossover filter and the subtractor and an adder coupled to the table and one of the gain adjustment devices, a method for processing audio data comprising:

- receiving audio input data for the associated audio input frequency band at the control signal processing system;

- generating the gain adjustment device control for the gain adjustment device;

- adjusting the gain adjustment device using the gain adjustment device control;

- decreasing the gain setting of the gain adjustment device for a predetermined period of time as the function of the transient in the associated audio input frequency band;

- generating a clipped audio input signal from the audio input data with a clipping system;

- generating high-pass filtered clipped audio data from the clipped audio input signal using a high pass filter coupled to the clipping system;

- multiplying the high-pass filtered clipped audio data by the predetermined value using the scaler multiplier coupled to the high pass filter;

receiving the difference signal at the table coupled to the scaler multiplier and selecting the predetermined value as the function of the difference signal;

generating the low frequency output and the high frequency output with a crossover filter coupled to the subtractor;

receiving the low frequency output and the high frequency output at the subtractor and generating the difference signal;

coupling the RMS converter between the crossover filter and the subtractor;

coupling the linear to log converter between the crossover filter and the subtractor; and

coupling the adder to the table and one of the gain adjustment devices.

20. A system for processing audio data comprising:

a first signal processing path configured to generate a mask control signal;

a second signal processing path configured to generate a decorrelated input audio signal; and

a mixer configured to mix the mask control signal and the decorrelated input audio signal and to generate an output.

21. The system of claim 20 wherein the first signal path further comprises a downward expander.

22. The system of claim 20 wherein the first signal path comprises:

a first automatic gain control unit configured to remove DC signal components from the input audio signal; and

a second automatic gain control unit coupled to the first automatic gain control unit and configured to normalize the DC compensated input audio signal with the input audio signal.

23. The system of claim 20 wherein the first signal path further comprises a de-interleaver.

24. The system of claim 20 wherein the first signal path further comprises a subtractor coupled to a maximum absolute value detector.

25. The system of claim 20 wherein the first signal path further comprises an automatic gain control multiplier coupled to the mixer.

26. The system of claim 20 wherein the second signal processing path further comprises a cross-over filter generating a low frequency output and a high frequency output.

27. The system of claim 26 wherein the second signal processing path further comprises a scaler coupled to the low frequency output.

28. The system of claim 26 wherein the second signal processing path further comprises an adder coupled to the low frequency output and the high frequency output.

29. The system of claim 27 wherein the second signal processing path further comprises a decorrelator coupled to the scaler.

30. A method for processing audio data comprising:
generating a mask control signal using a first signal processing path;
generating a decorrelated input audio signal using a second signal processing path; and
mixing the mask control signal and the decorrelated input audio signal with a mixer to generate an output.

31. The method of claim 30 wherein generating the mask control signal using the first signal processing path comprises generating a mask control signal using a downward expander.

32. The method of claim 30 wherein generating the mask control signal using the first signal processing path comprises:

removing DC signal components from the input audio signal using a first automatic gain control unit to generate a DC compensated input audio signal; and

normalizing the DC compensated input audio signal with the input audio signal using a second automatic gain control unit coupled to the first automatic gain control unit.

33. The method of claim 30 wherein generating the mask control signal using the first signal processing path comprises generating the mask control signal using a de-interleaver.

34. The method of claim 30 wherein generating the mask control signal using the first signal processing path comprises generating the mask control signal using a subtractor coupled to a maximum absolute value detector.

35. The method of claim 30 wherein generating the mask control signal using the first signal processing path comprises generating the mask control signal using an automatic gain control multiplier coupled to the mixer.

36. The method of claim 30 wherein generating a decorrelated input audio signal using a second signal processing path comprises generating a decorrelated input audio signal using a cross-over filter generating a low frequency output and a high frequency output.

37. The method of claim 36 wherein the second signal processing path further comprises a scaler coupled to the low frequency output.

38. The method of claim 36 wherein the second signal processing path further comprises an adder coupled to the low frequency output and the high frequency output.

39. In a system for processing audio data having a first signal processing path configured to generate a mask control signal, a second signal processing path configured to generate a decorrelated input audio signal, a mixer configured to mix the

mask control signal and the decorrelated input audio signal and to generate an output, wherein the first signal path further comprises a downward expander, a first automatic gain control unit configured to remove DC signal components from the input audio signal, a second automatic gain control unit coupled to the first automatic gain control unit and configured to normalize the DC compensated input audio signal with the input audio signal, a de-interleaver, a subtractor coupled to a maximum absolute value detector, an automatic gain control multiplier coupled to the mixer, a cross-over filter generating a low frequency output and a high frequency output, wherein the second signal processing path further comprises a scaler coupled to the low frequency output, an adder coupled to the low frequency output and the high frequency output, and a decorrelator coupled to the scaler, a method for processing audio data comprising:

- generating the mask control signal using the first signal processing path;

- generating the decorrelated input audio signal using the second signal processing path;

- mixing the mask control signal and the decorrelated input audio signal with the mixer to generate the output;

- wherein generating the mask control signal using the first signal processing path comprises:

 - generating the mask control signal using a downward expander;

 - removing DC signal components from the input audio signal using a first automatic gain control unit to generate a DC compensated input audio signal;

 - normalizing the DC compensated input audio signal with the input audio signal using a second automatic gain

control unit coupled to the first automatic gain control unit;

generating the mask control signal using a de-interleaver;

generating the mask control signal using a subtractor coupled to a maximum absolute value detector; and

generating the mask control signal using an automatic gain control multiplier coupled to the mixer; and

wherein generating a decorrelated input audio signal using a second signal processing path comprises:

generating a decorrelated input audio signal using a cross-over filter generating a low frequency output and a high frequency output;

wherein the second signal processing path further comprises a scaler coupled to the low frequency output; and

wherein the second signal processing path further comprises an adder coupled to the low frequency output and the high frequency output.

1/14

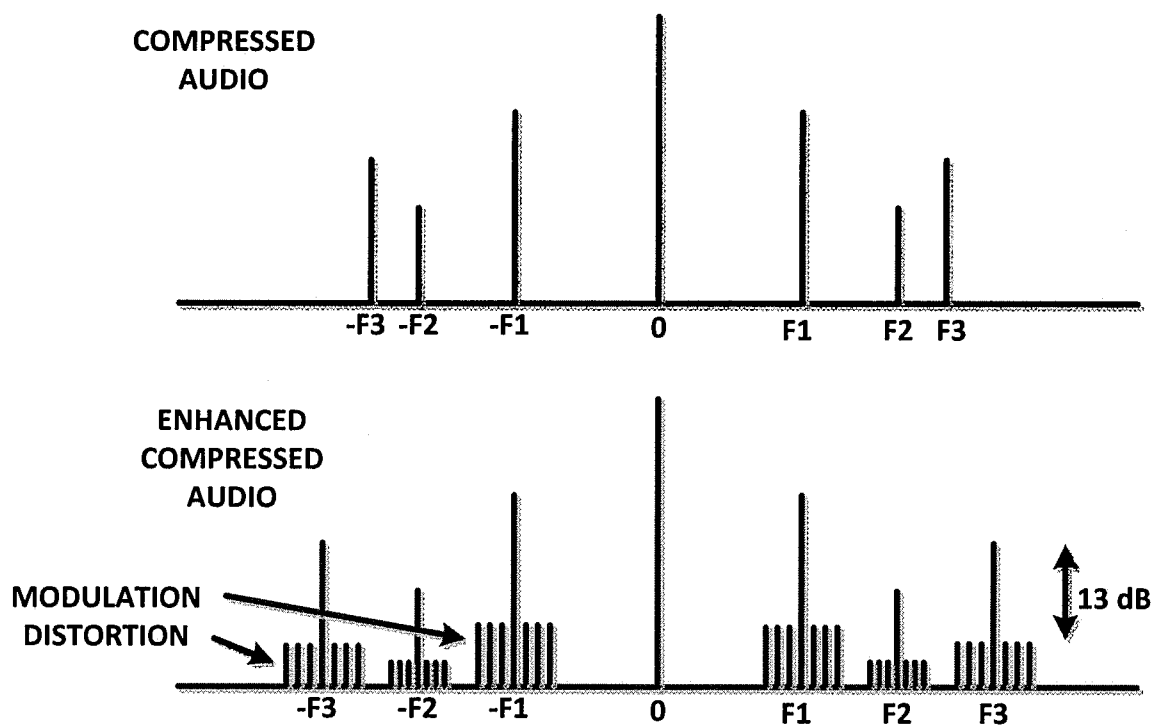


FIG. 1

100 ↑

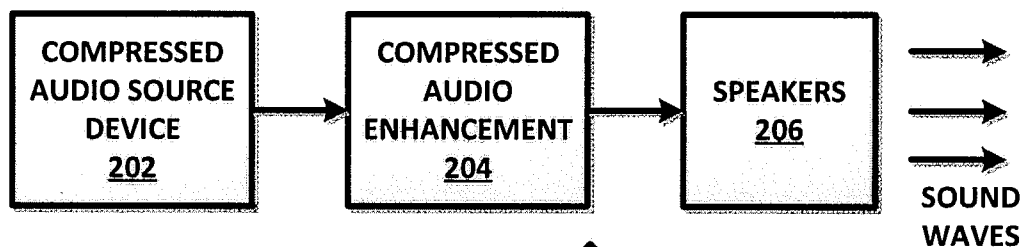


FIG. 2

200 ↑

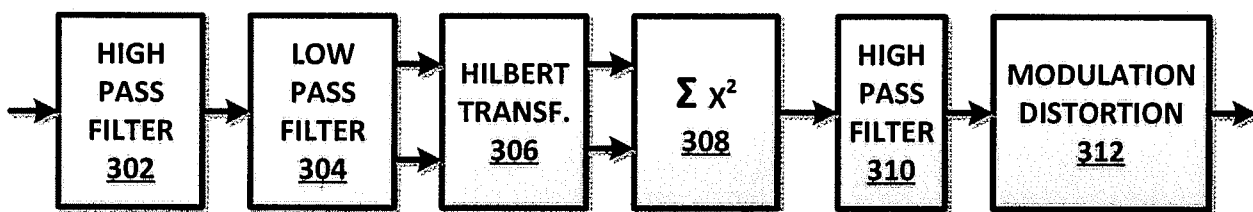


FIG. 3

300 ↑

2/14

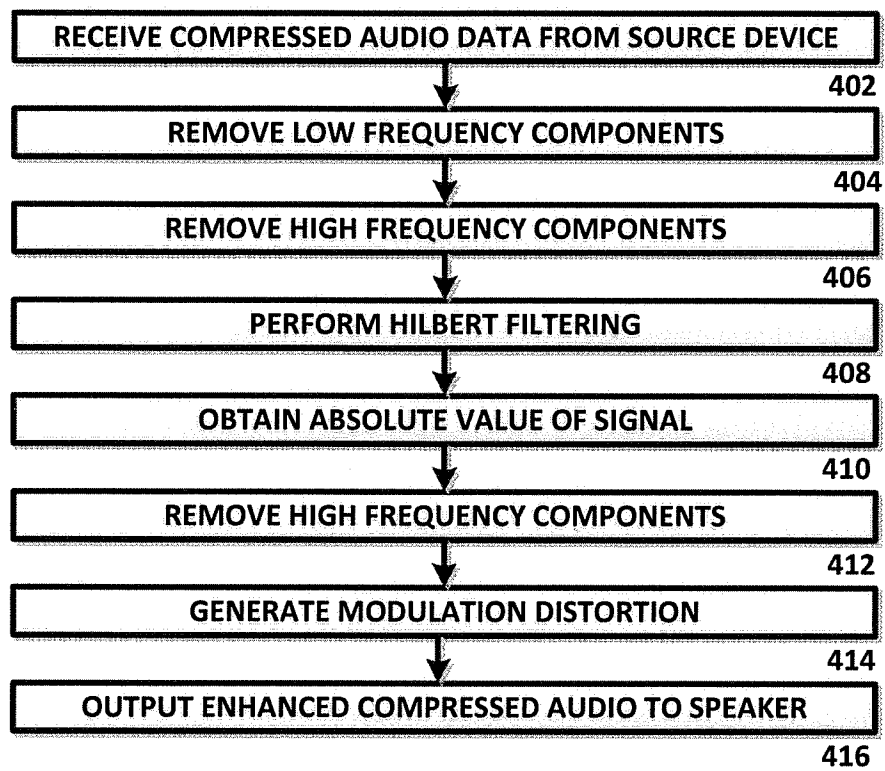


FIG. 4

400 ↑

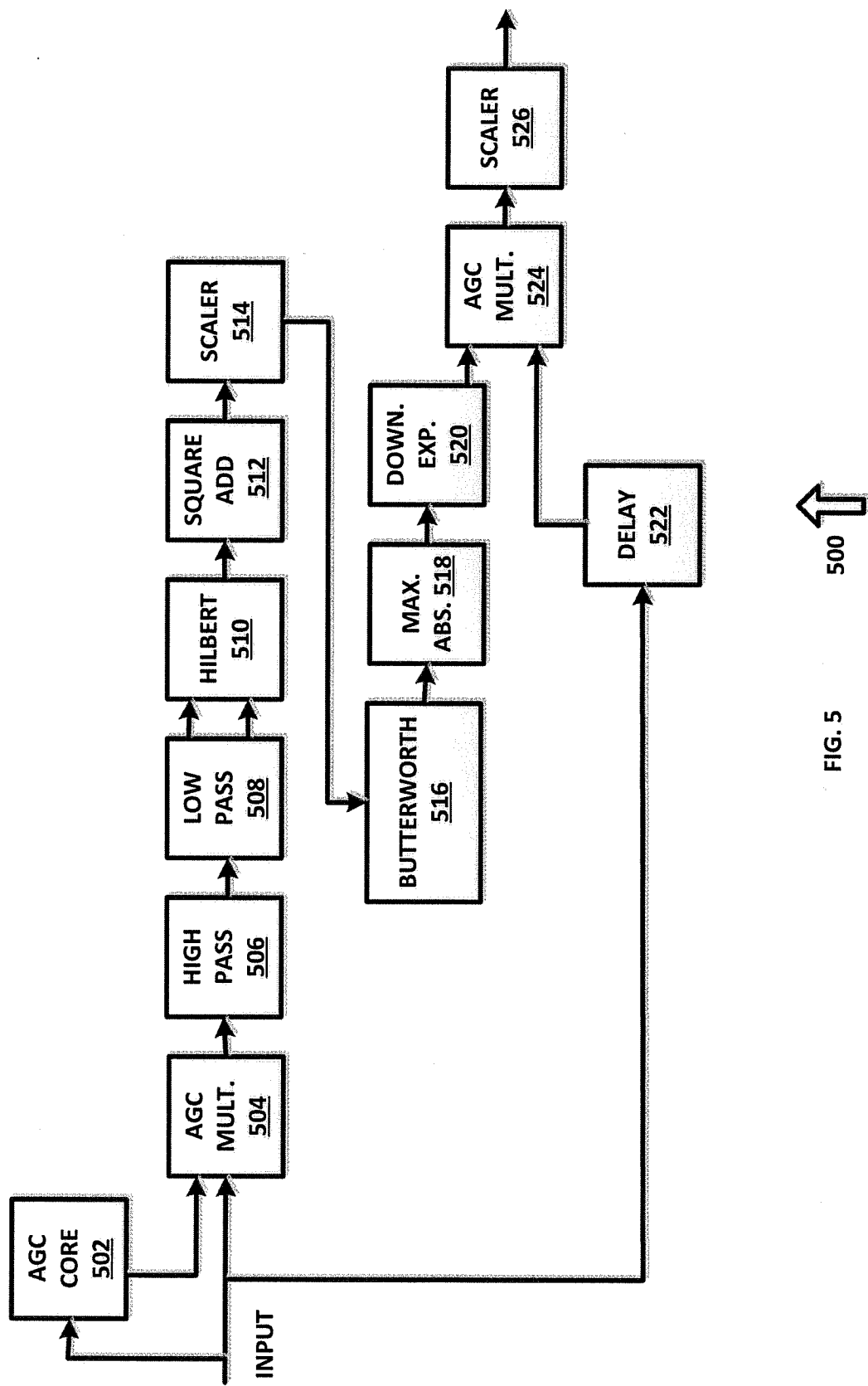


FIG. 5

4/14

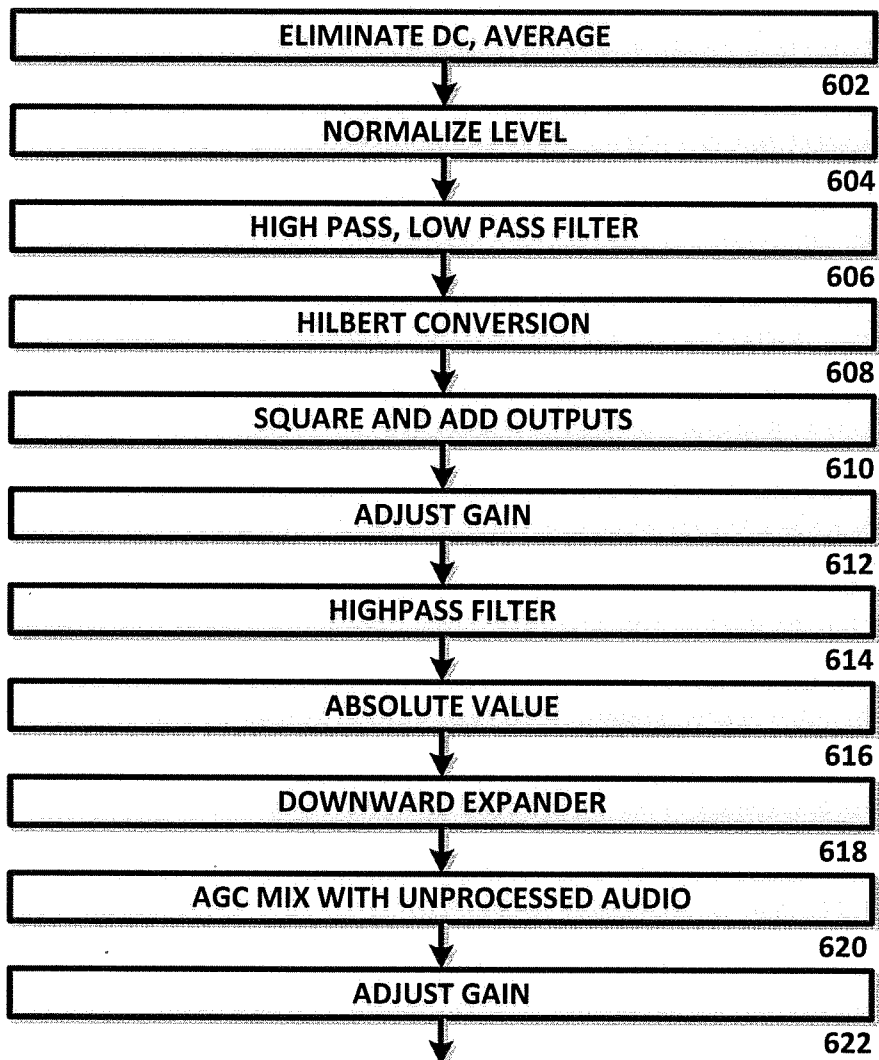


FIG. 6



5/14

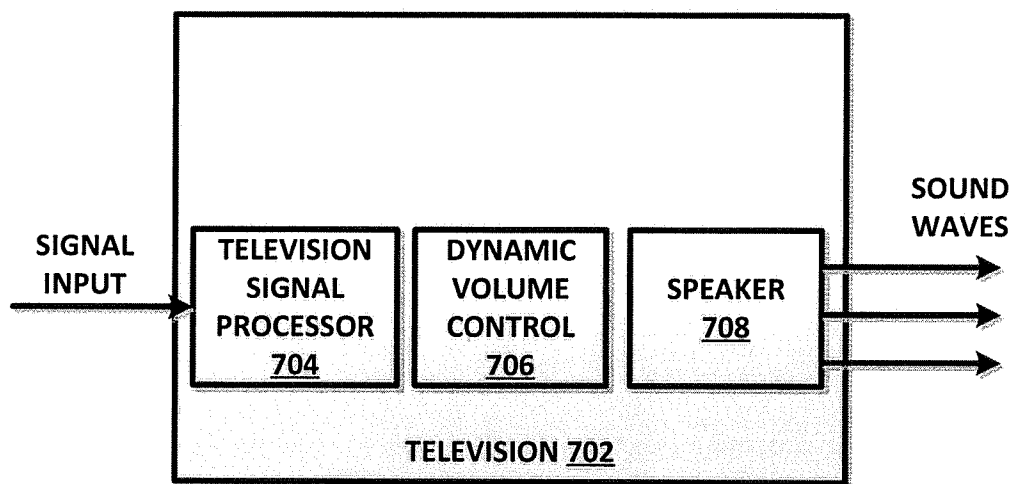


FIG. 7 700 ↑

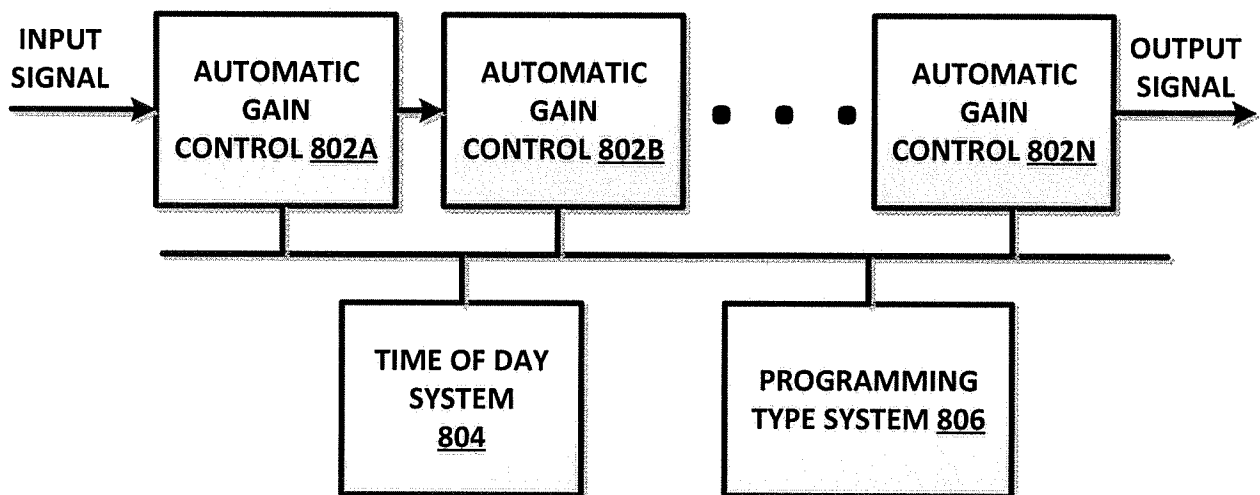


FIG. 8 800 ↑

6/14

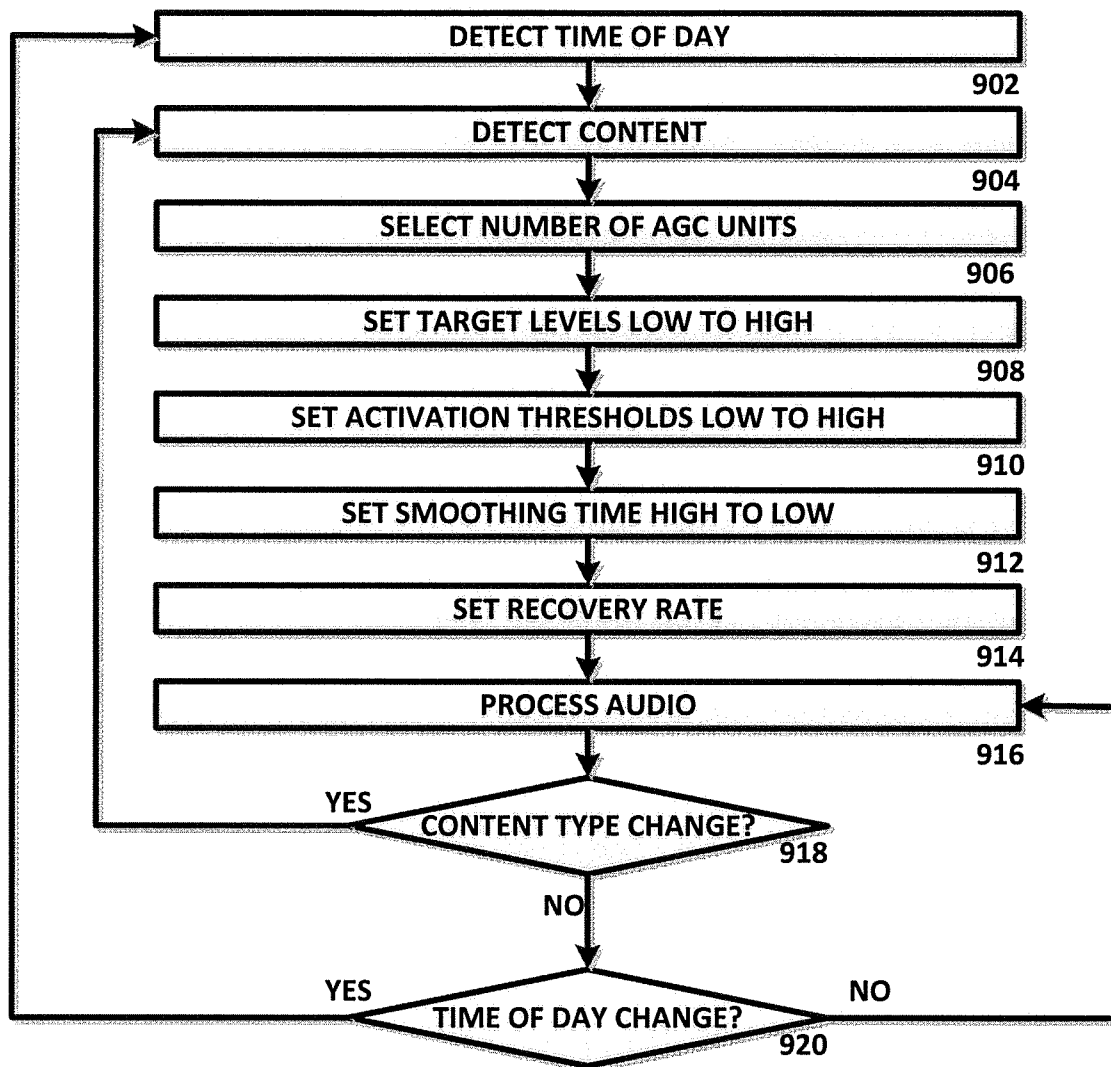
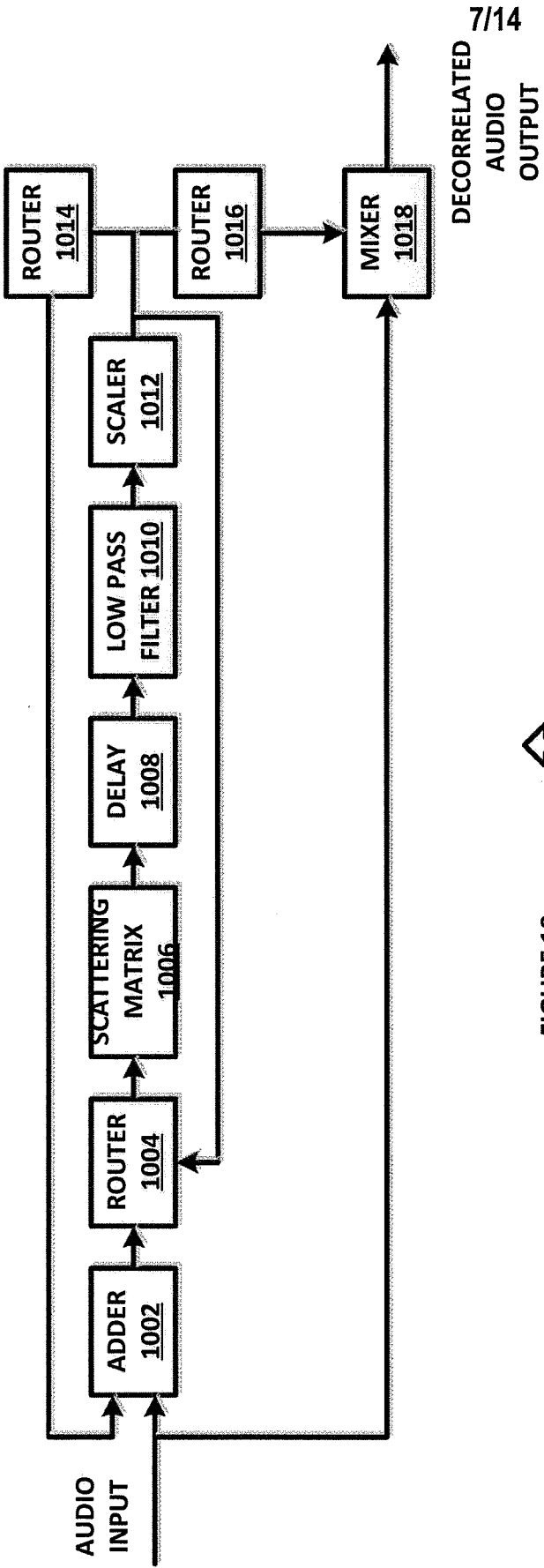


FIG. 9

900





8/14

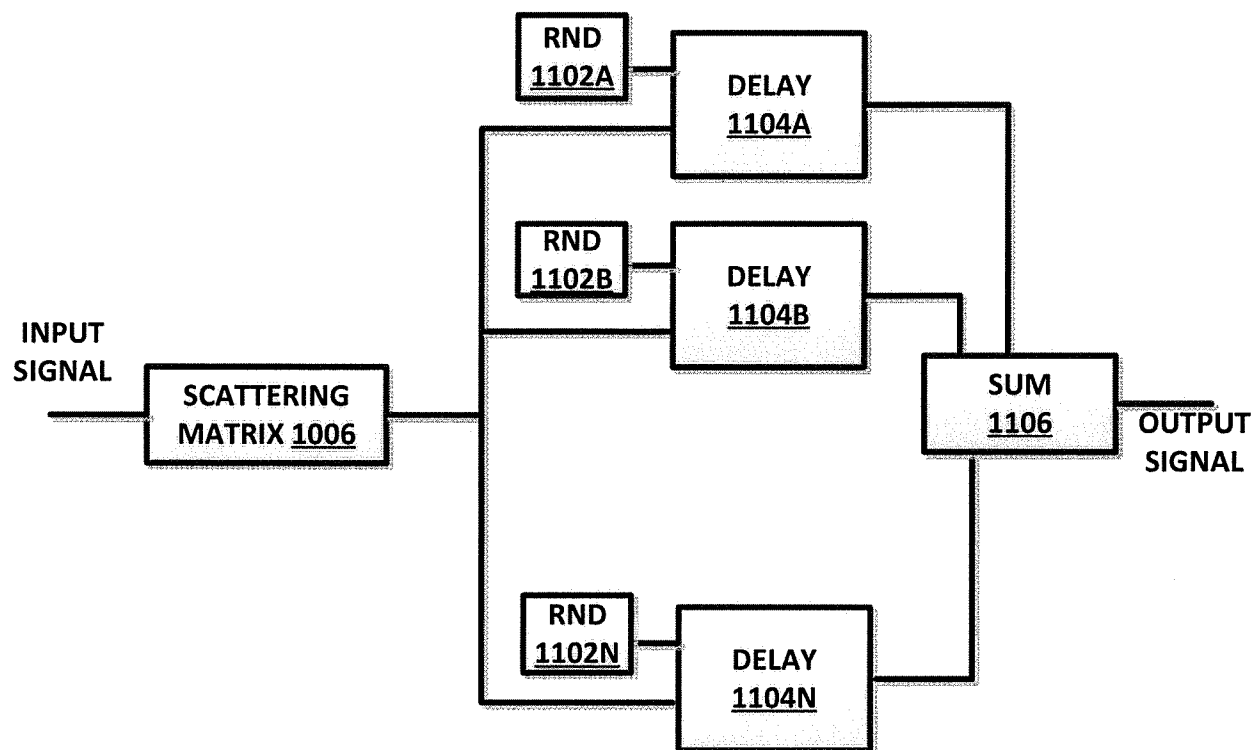


FIGURE 11

1100 ↑

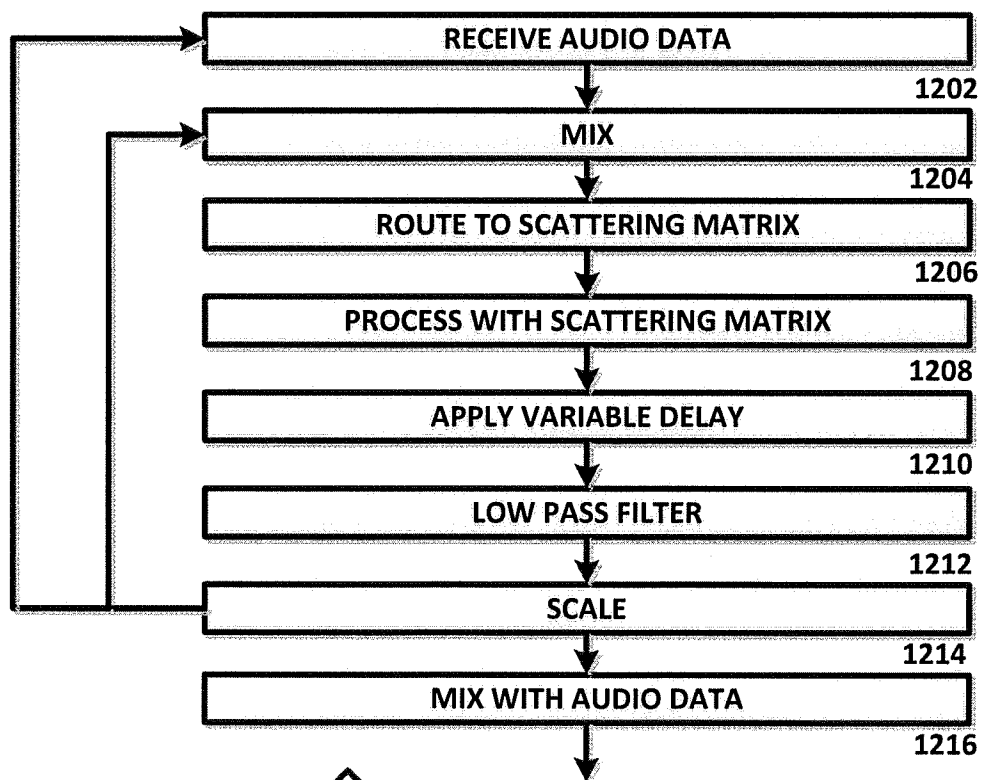


FIGURE 12

1200 ↑

9/14

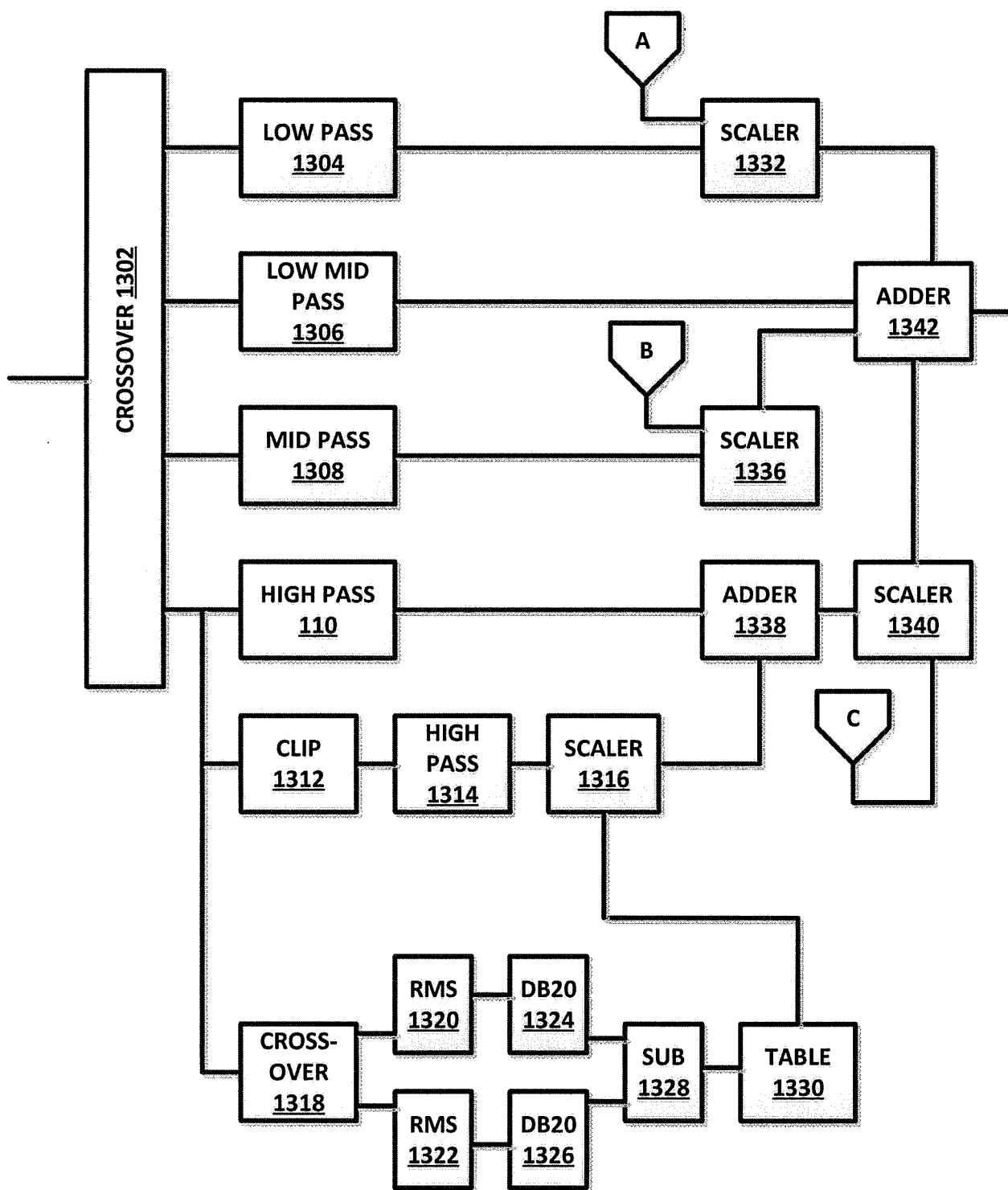


FIG. 13

1300



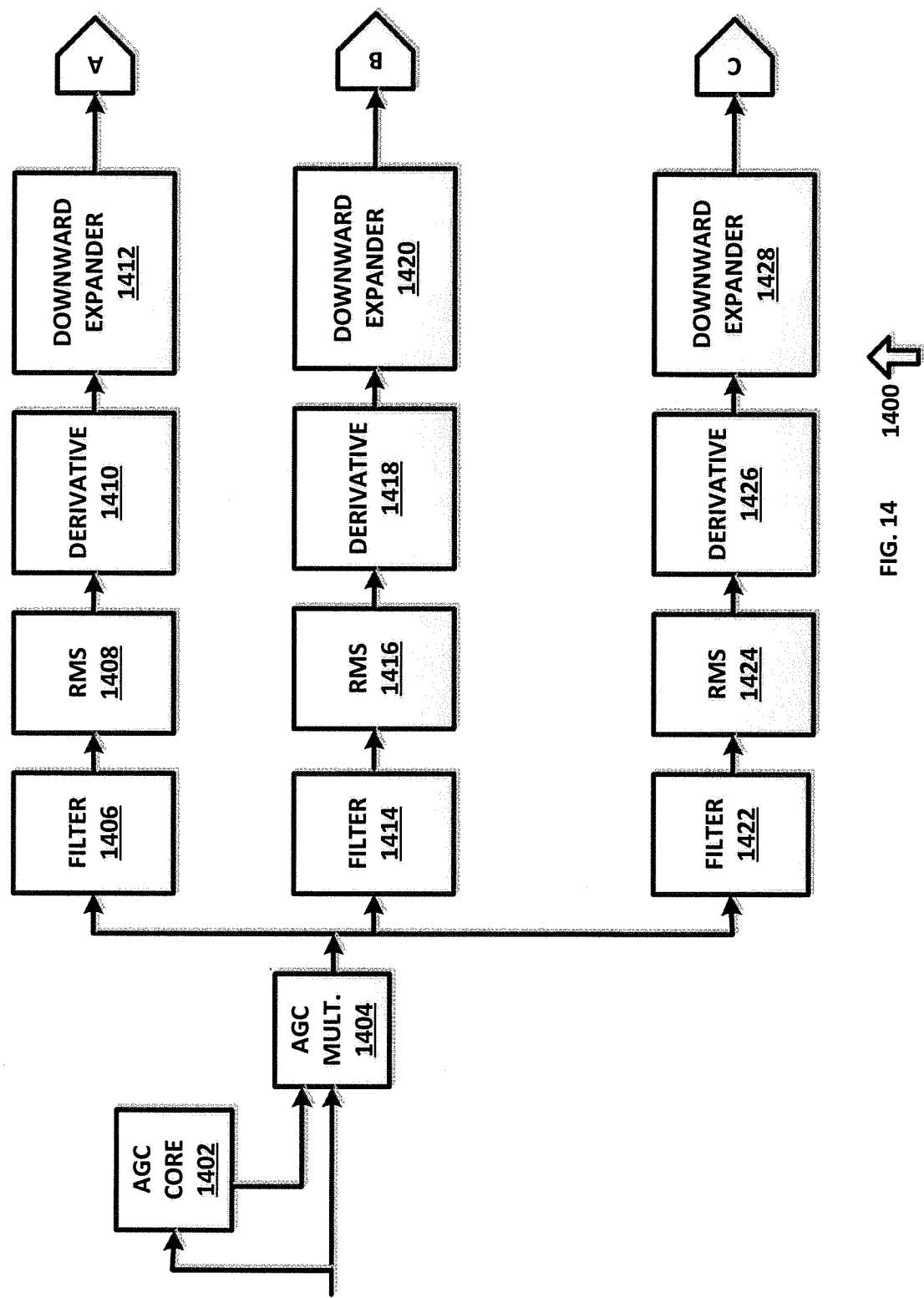


FIG. 14 1400

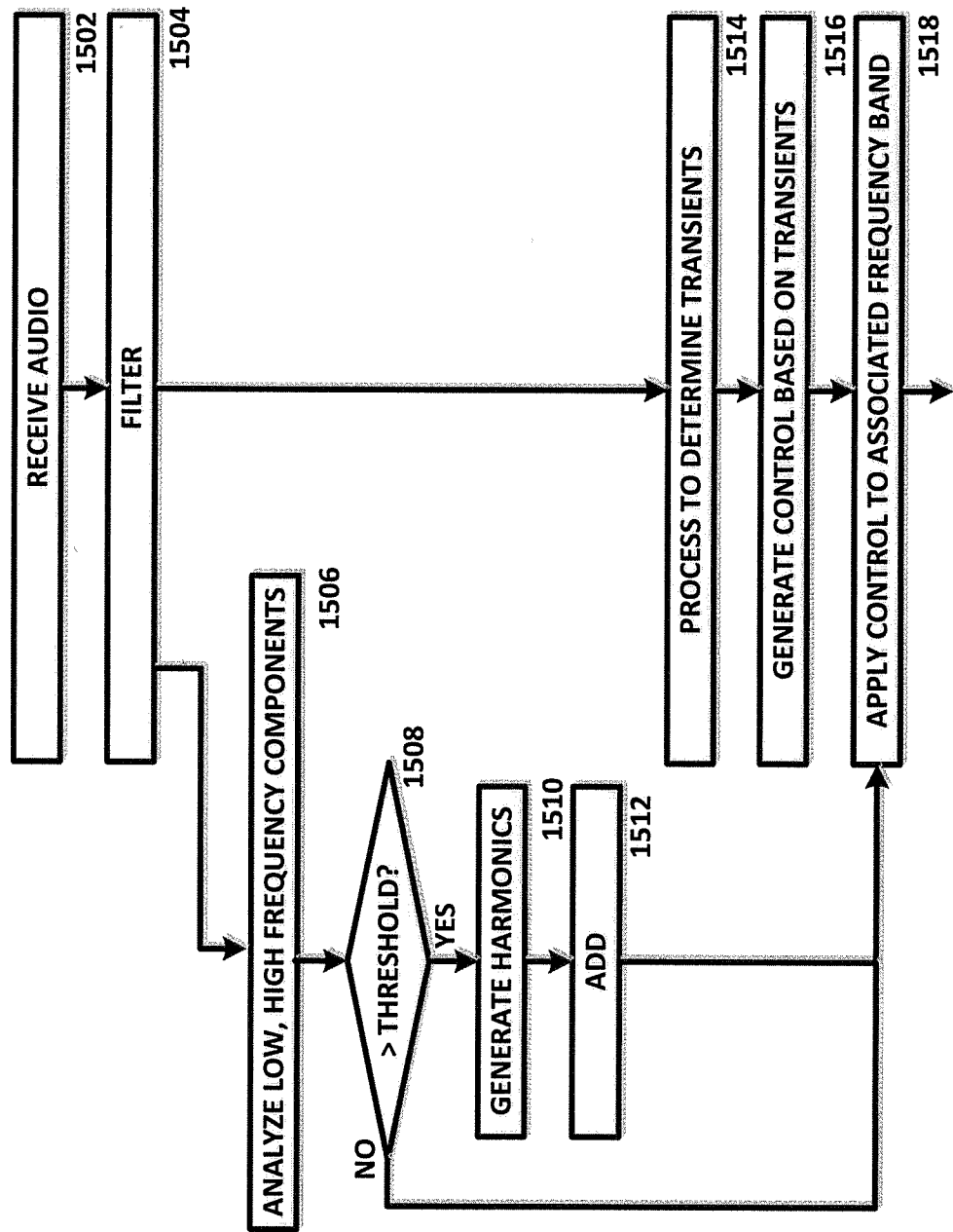
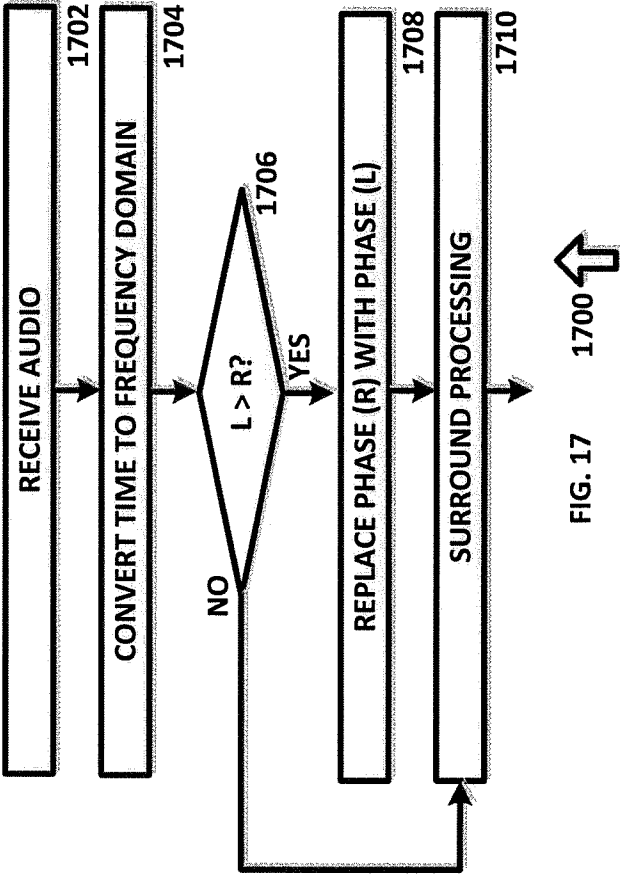
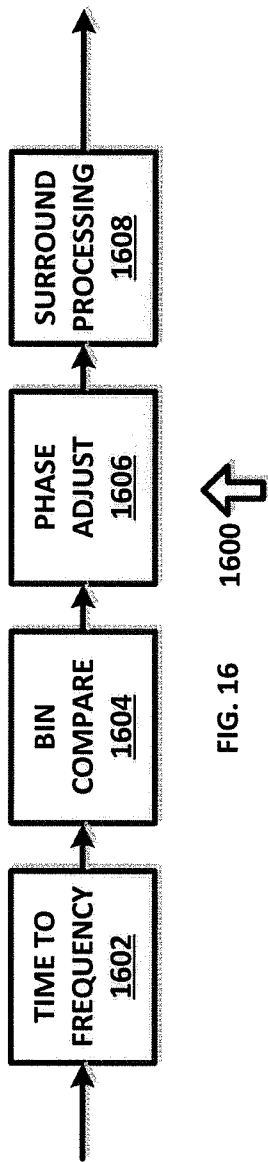


FIG. 15 1500



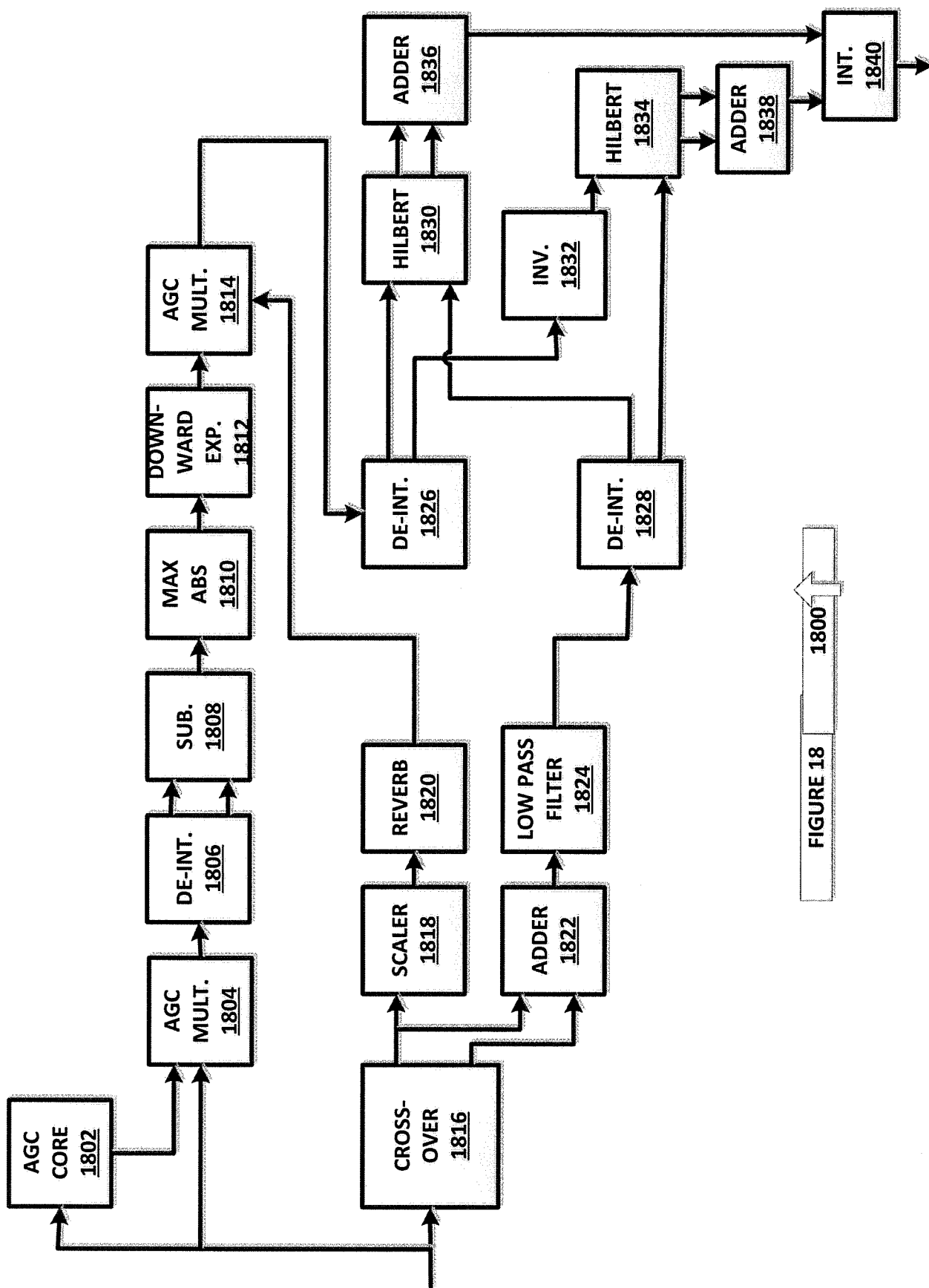


FIGURE 18 1800

14/14

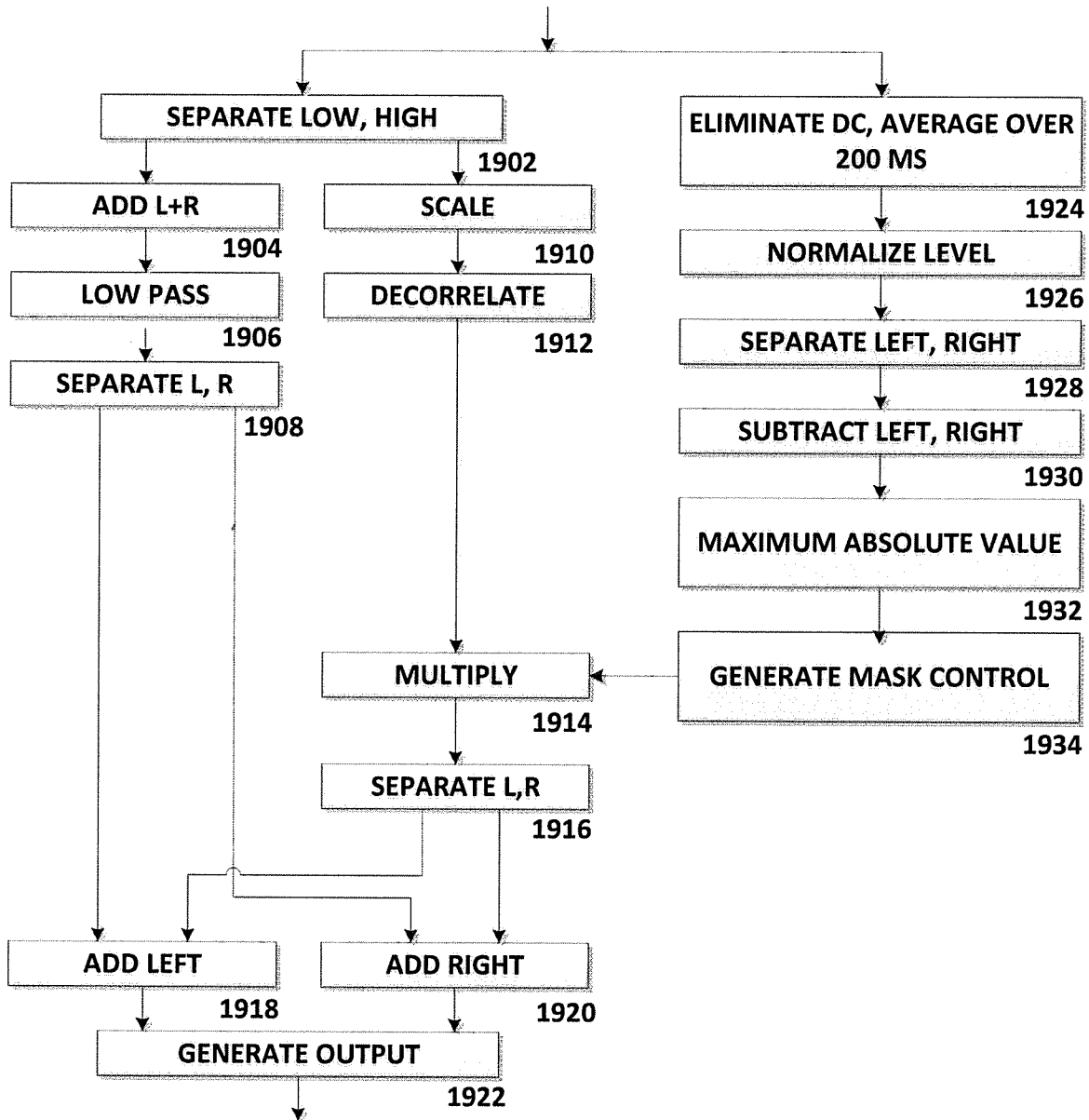


FIGURE 19

1900

