SYSTEM AND METHOD FOR COMPRESSED AUDIO ENHANCEMENT

Applicant: Psyx Research, Inc., Irvine, CA (US)

Inventor: Robert W. Reams, Bellevue, WA (US)

Appl. No.: 14/863,368

Filed: Sep. 23, 2015

Related U.S. Application Data

Provisional application No. 62/092,603, filed on Dec. 16, 2014, provisional application No. 62/133,167, filed on Mar. 13, 2015, provisional application No. 62/156,061, filed on May 1, 2015, provisional application No. 62/156,065, filed on May 1, 2015.

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 listener active. One or more speakers configured to receive the enhanced audio data and to generate sound waves using the enhanced audio data.
COMPRESSED AUDIO

ENHANCED COMPRESSED AUDIO

MODULATION DISTORTION

FIG. 1

COMPRESSED AUDIO SOURCE DEVICE 202

COMPRESSED AUDIO ENHANCEMENT 204

SPEAKERS 206

FIG. 2

HIGH PASS FILTER 302

LOW PASS FILTER 304

HILBERT TRANSF. 306

\( \sum x^2 \)

HIGH PASS FILTER 310

MODULATION DISTORTION 312

FIG. 3
RECEIVE COMPRESSED AUDIO DATA FROM SOURCE DEVICE

REMOVE LOW FREQUENCY COMPONENTS

REMOVE HIGH FREQUENCY COMPONENTS

PERFORM HILBERT FILTERING

OBTAIN ABSOLUTE VALUE OF SIGNAL

REMOVE HIGH FREQUENCY COMPONENTS

GENERATE MODULATION DISTORTION

OUTPUT ENHANCED COMPRESSED AUDIO TO SPEAKER
SYSTEM AND METHOD FOR COMPRESSED AUDIO ENHANCEMENT

RELATED APPLICATIONS


TECHNICAL FIELD

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

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

SUMMARY OF THE INVENTION

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.

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

Aspects of the disclosure can be better understood with reference to the following drawings. The components in the drawings 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:

FIG. 1 is a diagram of a frequency diagram showing the effect of compressed audio processing in accordance with the present disclosure;

FIG. 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;

FIG. 3 is a diagram of a system for providing modulation distortion in accordance with an exemplary embodiment of the present disclosure; and

FIG. 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.

DETAILED DESCRIPTION OF THE INVENTION

In the description that follows, like parts are marked throughout the specification and drawings with the same reference numerals. The drawing figures might not 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.

FIG. 1 is a diagram of a frequency diagram 100 showing the effect of compressed audio processing in accordance with the present disclosure. Frequency diagram 100 shows a frequency distribution for compressed audio data with frequency components at +/- F1, F2 and F3. These frequency components are relatively sparse. Frequency diagram 100 also shows a frequency distribution for enhanced compressed audio data with frequency components centered at +/- F1, F2 and F3 and associated modulation distortion components in a range around the centered frequency components. Although the modulation distortion 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 is also at least 13 dB below the signal magnitude, in order to mask the modulation distortion components from perception by the user.

Typically, modulation distortion 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, 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 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.

FIG. 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.

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 a microcomputer or other suitable controller, 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.

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, cellular telephone, an automobile audio system or other suitable audio systems that are configured to provide streaming compressed audio data.

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 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.

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, headset speakers, cellular telephone speakers, sound bar speakers, computer speakers or other suitable speakers.

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.
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 can be provided to an amplifier, a speaker or other suitable devices.

[0027] In operation, system 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 state 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.

[0028] FIG. 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. Algorithm 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.

[0029] Algorithm 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.

[0030] 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.

[0031] 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.

[0032] 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.

[0033] 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.

[0034] 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.

[0035] 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 0.1616 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.

[0036] 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.

[0037] 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 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; and
one or more speakers configured to receive the enhanced audio data and to generate sound waves using the enhanced audio data.

2. The system of claim 1 wherein the compressed audio enhancement system comprises 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.

3. The system of claim 1 wherein the compressed audio enhancement system comprises 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.

4. The system of claim 1 wherein the compressed audio enhancement system comprises 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.

5. The system of claim 1 wherein the compressed audio enhancement system comprises 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.

6. The system of claim 1 wherein generation of the enhanced audio data comprises 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.

7. The system of claim 1 wherein generation of the enhanced audio data comprises 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.

8. The system of claim 1 wherein generation of the enhanced audio data comprises 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.

9. The system of claim 1 wherein the compressed audio enhancement system comprises a downward expander.

10. The system of claim 1 wherein the compressed audio enhancement system comprises a downward expander having an attack time of less than one millisecond.

11. A method for processing digitally-encoded audio data 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; and

generating sound waves with a sound wave generating device using the modified digitally encoded audio data.

12. The method of claim 11 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.

13. The method of claim 11 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.

14. The method of claim 11 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.

15. The method of claim 11 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.

16. The method of claim 11 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.

17. The method of claim 11 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.

18. The method of claim 11 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.

19. The method of claim 11 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.

20. 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 the 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, 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 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 one millisecond; and
- 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.

* * * * *