



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
01.06.2022 Bulletin 2022/22

(51) International Patent Classification (IPC):
H04S 1/00 (2006.01) H04S 5/00 (2006.01)
H04S 7/00 (2006.01)

(21) Application number: **20210629.0**

(52) Cooperative Patent Classification (CPC):
H04S 5/00; H04S 1/002; H04S 1/007; H04S 7/305

(22) Date of filing: **30.11.2020**

(84) Designated Contracting States:
**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO
PL PT RO RS SE SI SK SM TR**
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(71) Applicant: **ASK Industries GmbH**
94559 Niederwinkling (DE)

(72) Inventor: **PAPOULIS, Eftychios**
94559 Niederwinkling (DE)

(74) Representative: **Hafner & Kohl PartmbB**
Schleiermacherstraße 25
90491 Nürnberg (DE)

(54) **METHOD OF PROCESSING AN INPUT AUDIO SIGNAL FOR GENERATING A STEREO OUTPUT AUDIO SIGNAL HAVING SPECIFIC REVERBERATION CHARACTERISTICS**

(57) A method of processing an input audio signal for generating a stereo output audio signal with a specific reverberation, the method comprising the following steps:

a) providing an input audio signal, the input audio signal being a mono input audio signal or a stereo input audio signal;

b) providing pre-recorded stereo Room-Impulse-Response (RIR) data of a specific acoustic environment, the RIR data comprising a defined number of RIR samples, the RIR data comprising an equal number of left channel samples and right channel samples;

c) determining a first number of RIR samples representing a stereo part of the RIR data and a second number of RIR samples representing a mono part of the RIR data, whereby the stereo part of the RIR data comprises a number of left channel samples for the left output channel and an equal number of right channel samples for the right output channel, and whereby the mono part of the RIR comprises a number of samples to be used for both the left and the right output channel;

d) subdividing the samples of the RIR into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono part of the RIR, whereby the duration that corresponds to the stereo part of the RIR and the duration that corresponds to the mono part of the RIR add up to the total duration of the RIR;

e1-e2) applying a first signal processing rule consisting in (e1) convolving the input audio signal with the left channel samples of the stereo part of the RIR data and (e2)

convolving the input audio signal with the right channel samples of the stereo part of the RIR data, thereby obtaining a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of input audio signal from the first group of samples of the RIR;

e3) applying a second signal processing rule consisting in convolving the mono input audio signal, or the mono version of the stereo input audio signal, with the mono part of the RIR data, thereby obtaining a processed mono audio signal part representing the reverberation of the input audio signal from the second group of samples of the RIR; and

f1) mixing the left channel audio signal part resulting from the processing of the input audio signal with the left channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a left channel output signal; and

f2) mixing the right channel audio signal part resulting from the processing of the input audio signal with the right channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a right channel output signal.

HUPOLS (Hybrid Mono-Stereo Uniform Partition Overlap-Save) implementation aspects of the above-mentioned signal processing method are provided. HUPOLS method is based on the UPOLS (Uniform Partition Overlap-Save) method.

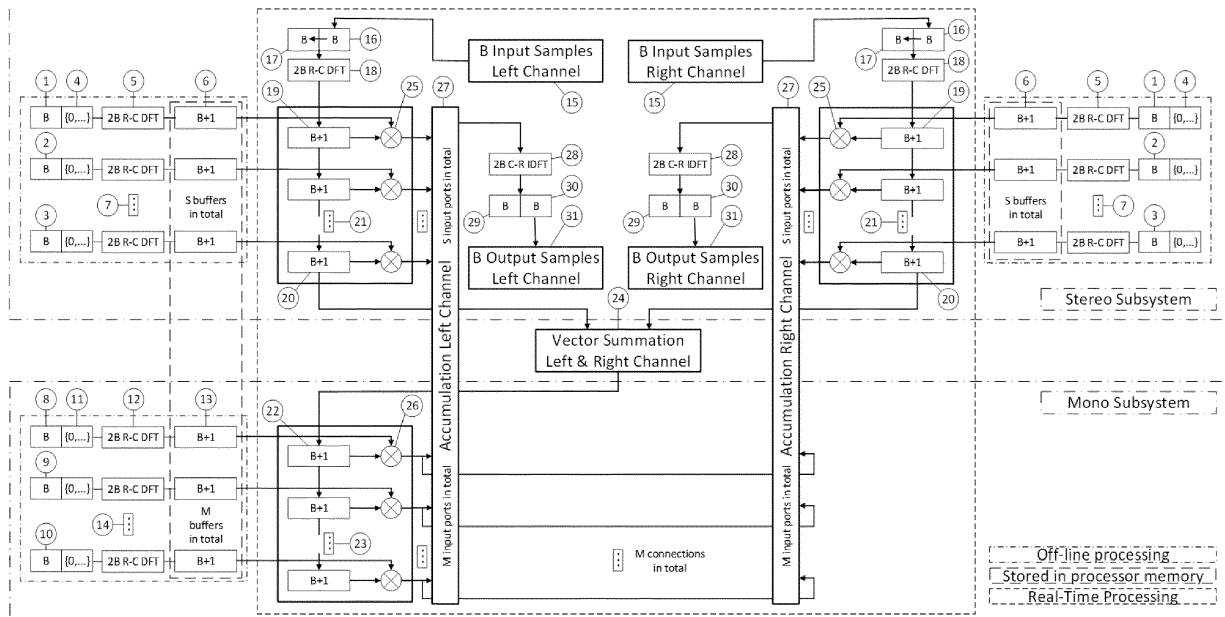


Fig. 1

Description

[0001] The invention relates to a method of processing an input audio signal for generating a stereo output audio signal having specific reverberation characteristics.

[0002] Audio signal processing generally, comprises processing of input audio signals, i.e. audio signals which are input to a digital signal processing unit ("DSP unit"), having specific audio signal properties, so as to generate output audio signals, i.e. audio signals which are output of the audio signal processing unit, having specific audio signal properties, at least partly different from the input audio signal properties. Specifically, audio signal processing may comprise modifying one or more properties of an input audio signal so as to obtain an output audio signal having one or more properties which are modified relative to the respective properties of the input audio signal.

[0003] A specific aim in audio signal processing comprises processing of an input audio signal for generating an output audio signal having specific reverberation characteristics, such as the specific reverberation characteristics of a specific acoustic environment, e.g. a specific room or venue.

[0004] In particular, it is known that a stereophonic real-time convolution-based artificial reverberation of a stereo input audio signal, using pre-recorded stereo room impulse response ("RIR") data from a real-life room, often requires more computing operations, such as Floating-Point Operations Per Input Sample ("FLOPIS"), and memory than available in a common DSP-unit, such as in a common DSP-unit of a vehicle.

[0005] The required number of FLOPIS and memory size typically, depend on the size of the room and the sampling rate used during the recording of the RIR data and the play-back of the audio signal to be reverberated.

[0006] For the sampling rates typically used for audio, and for the reverberation times of large rooms, the length of the RIR data turns out to be very large. As an example, for the sampling rate of 48 KHz and for the reverberation time of 2 seconds (that the interior of a large room can exhibit), the stereo RIR model has $2 \times 96 \times 10^3$ samples. A direct convolution in the time-domain would require $2 \times 192 \times 10^3$ FLOPIS and $2 \times 192 \times 10^3$ memory locations to store the $2 \times 96 \times 10^3$ RIR samples plus the $2 \times 92 \times 10^3$ most recent samples of the input signal. In these figures, the factor of two accounts for the stereo property of the audio signals. Also, the figure of 192×10^3 comes from the fact that each multiply-and-add operation counts as two FLOPIS.

[0007] These numbers are very large, even for modern hardware. Moreover, they imply a very large memory-throughput, namely, number of memory-access operations per input audio sample.

[0008] Hence, there exists a need for improved methods for processing an input audio signal for generating a stereo output audio signal having specific reverberation characteristics, i.e. particularly (the) specific reverbera-

tion characteristics of a specific acoustic environment, particularly with respect to the computing power and memory requirements for the respective DSP-unit.

[0009] In fact, diverse approaches for real-time artificial reverberation are known. As an example, the Uniform Partition Overlap-Save ("UPOLS") method is a widely used uniform partition algorithm for real-time artificial reverberation. UPOLS may significantly reduce the computing operations compared with the direct convolution. However, UPOLS doubles the required memory because UPOLS works with complex data.

[0010] It is thus, the object of the present invention to provide an improved method of processing an input audio signal for generating a stereo output audio signal with specific reverberation characteristics, particularly specific reverberation characteristics of a specific acoustic environment, particularly with respect to the computing power and memory requirements for digital audio signal processing and with respect to the ease of implementation.

[0011] The object is achieved by a method of processing an input audio signal for generating a stereo output audio signal having specific reverberation characteristics according to Claim 1. The Claims depending on Claim 1 refer to possible embodiments of the method of Claim 1.

[0012] A first aspect of the invention refers to a method of processing an input audio signal for generating a stereo output audio signal with specific reverberation characteristics. The method thus, enables generating a stereo output audio signal having specific reverberation characteristics, i.e. specific reverberation characteristics, such as the specific reverberation characteristics of a specific acoustic environment, e.g. a specific room of a specific building, by processing of an input audio signal.

[0013] The method can be implemented by a hardware- and/or software-embodied digital signal processing unit ("DSP-unit") which is configured to perform the method. The DSP unit may comprise at least one processing unit, such as a processor, and at least one memory unit. The DSP unit may form part of an apparatus for processing an input audio signal so as to generate a stereo output audio signal. A respective apparatus can form a vehicle audio system or a car audio system, i.e. an audio system that is to be installed or is installed in a vehicle or a car, respectively. Alternatively, a respective apparatus can form part of a respective vehicle audio system or car audio system, respectively.

[0014] The basic steps of the method of processing an input audio signal for generating a stereo output audio signal with specific reverberation characteristics will be specified in the following:

According to a first step of the method, an input audio signal is provided. The input audio signal is a mono input audio signal ("mono signal") or a stereo input audio signal ("stereo signal"). As will be apparent from the further specification, the input audio signal is the signal which is to be processed in accordance with the method so as to generate a stereo output audio signal of a specific rever-

beration. The generated output audio signal is always a stereo audio signal, regardless of whether the input audio signal is mono or stereo.

[0015] The first step of the method can be implemented by a hardware- and/or software-embodied input audio signal providing unit which is configured to provide an input audio signal that is mono or stereo. An input audio signal can be a music- and/or speech-signal, i.e. a signal comprising music- and/or speech-content.

[0016] According to a second step of the method, pre-recorded stereo data of the Room-Impulse-Response ("RIR") of a specific acoustic environment, such as a specific building or venue, are provided. The pre-recorded RIR data, respectively comprise a defined number of RIR samples, in particular left channel samples and right channel samples. The RIR data are or comprise thus, typically stereo data. Particularly, the RIR data comprise an equal number of left channel samples and right channel samples. The RIR data can be obtained through known methods for recording the RIR of acoustic environments. The actual recording of a respective acoustic environment is typically, not a step of the method.

[0017] The second step of the method can be implemented by a hardware- and/or software-embodied RIR data providing unit which is configured to provide pre-recorded RIR data of a specific acoustic environment.

[0018] According to a third step of the method, a first number of RIR samples representing a stereo part of the RIR represented by the RIR data and a second number of RIR samples representing a mono part of the RIR represented by the RIR data is determined. Thereby, the stereo part of the RIR comprises a number of left channel RIR samples for the left output channel and an equal number of right channel RIR samples for the right output channel. The mono part of the RIR represented by the RIR data comprises a number of RIR samples to be used for both the left and the right output channel. The third step of the method thus, comprises the determination of a stereo part of the RIR which is or can be represented by a first number of RIR samples and a mono part of the RIR which is or can be represented by a second number of RIR samples. The stereo part of the RIR is the first part of the RIR and is followed by the mono part of the RIR which is the second part of the RIR. The duration of the stereo part and of the mono part of the RIR typically, add up to the total duration of the RIR.

[0019] The third step of the method can be implemented by a hardware- and/or software-embodied determining unit which is configured to determine a first number of RIR samples representing a stereo part of the RIR and a second number of RIR samples representing a mono part of the RIR, whereby the stereo part of the RIR comprises a number of left channel RIR samples for the left output channel and an equal number of right channel RIR samples for the right output channel. The mono part of the RIR comprises a number of RIR samples for both the left and the right output channel.

[0020] According to a fourth step of the method, the

samples of the RIR are subdivided into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono part of the RIR. Hence, the first group of RIR samples typically comprises the first number of RIR samples representing the stereo part of the RIR and the second group of RIR samples typically comprises the second number of RIR samples representing the mono part of the RIR. The first group of RIR samples typically, represents a (distinct) early reflections part ("ERP") of the RIR and the second group of RIR samples typically, represents a (distinct) late reflections part ("LRP") of the RIR. The first group of RIR samples can comprise a period ranging between 1 ms and 150 ms, particularly between 10 ms and 100 ms, of the initial duration of the RIR. The second group of RIR samples comprises the remaining duration of the RIR.

[0021] The fourth step of the method can be implemented by a hardware- and/or software-embodied subdividing unit which is configured to subdivide the samples of the RIR into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono part of the RIR.

[0022] The third and fourth step of the method can be combined in one step which comprises both the determining aspect as specified above in context with the third step and the subdividing aspect as specified above in context with the fourth step. Hence, a hardware- and/or software-embodied determining and subdividing unit which is configured to determine a first number of RIR samples representing a stereo part of the RIR and a second number of RIR samples representing a mono part of the RIR, whereby the stereo part of the RIR comprises a number of left channel RIR samples for the left output channel and an equal number of right channel RIR samples for the right output channel and configured to subdivide the samples of the RIR into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono part of the RIR, can be used when the third and fourth step are combined.

[0023] According to a fifth step of the method, a first signal processing rule for processing, particularly by convolving, the input audio signal with the left channel samples of the stereo part of the RIR and for processing, particularly by convolving, the input audio signal with the right channel samples of the stereo part of the RIR is applied or implemented. Thereby, a processed left channel audio signal part and a processed right channel audio signal part, representing the reverberation of the input audio signal - which can be a mono input audio signal or a stereo input audio signal - from the first group of samples of the RIR is obtained. Hence, by applying or implementing the first signal processing rule, the input audio signal is processed, i.e. typically convolved, with the left channel samples of the stereo part of the RIR, whereby a processed, i.e. typically convolved, left channel audio signal part is obtained. Likewise, the input audio signal

is processed, i.e. typically convolved, with the right channel samples of the stereo part of the RIR, whereby a processed, i.e. typically convolved, right channel audio signal part is obtained. The respective processed left and right channel audio signal parts represent the reverberation, i.e. specifically the artificially generated reverberation, of the input audio signal from the first group of samples of the RIR.

[0024] Further, the fifth step comprises applying or implementing a second signal processing rule for processing, particularly by convolving, the input audio signal with the mono part of the RIR data. Thereby, if the input audio signal is a mono input audio signal, a processed mono audio signal part representing the reverberation of the mono input audio signal from the second group of samples of the RIR is obtained; and, if the input audio signal is a stereo input audio signal, a processed mono audio signal part representing the reverberation of the mono version of the stereo input audio signal from the second group of samples of the RIR is obtained. Hence, by applying or implementing the second signal processing rule, a mono input audio signal is processed so as to obtain a processed mono audio signal part representing the reverberation of the mono input audio signal from the second group of samples of the RIR (the mono part of the RIR), and a stereo input audio signal is processed so as to obtain a processed mono audio signal part representing the reverberation of both the left and right channel of the stereo input audio signal from the second group of samples of the RIR (the mono part of the RIR). The signal resulting from the processing of the input audio signal with the mono part of the RIR, namely with the second group of samples of the RIR, is always a mono signal, regardless if the input audio signal is a mono or a stereo signal. Thus, the mono input audio signal, or the stereo input audio signal after being converted into a mono input audio signal, is processed with the second group of samples of the RIR, namely with the mono part of the RIR, to generate a processed mono audio signal, representing the reverberation of the input audio signal with the second group of samples of the RIR, namely with the mono part of the RIR.

[0025] The fifth step of the method can be implemented by a hardware- and/or software-embodied signal processing structure or unit which is configured to apply or implement a first signal processing rule for processing, particularly by convolving, the input audio signal with the left channel samples of the stereo part of the RIR data and for processing, particularly by convolving, the input audio signal with the right channel samples of the stereo part of the RIR data, thereby obtaining a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of the mono or stereo input audio signal from the first group of samples of the RIR. Further, the hardware- and/or software-embodied signal processing structure or unit is configured to apply or implement a second signal processing rule for processing, particularly by convolving, the input

audio signal with the mono part of the RIR data, thereby obtaining, a processed mono audio signal part representing the reverberation of the mono or stereo input audio signal from the second group of samples of the RIR.

[0026] According to a sixth step of the method, the left channel audio signal part resulting from the processing of the input audio signal with the left channel samples of the stereo part of the RIR is mixed with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, whereby a left channel output signal is generated. Further, the right channel audio signal part resulting from the processing of the input audio signal with the right channel samples of the stereo part of the RIR is mixed with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, whereby a right channel output signal is generated. Hence, by mixing the left channel audio signal part with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, a left channel output signal is generated; and by mixing the right channel audio signal part with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, a right channel output signal is generated. The generated left and right channel output signals build the stereo output audio signal having the specific reverberation characteristics.

[0027] The sixth step of the method can be implemented by a hardware- and/or software-embodied mixing unit which is configured to mix the left channel audio signal part resulting from the processing of the input audio signal with the left channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a left channel output signal; and to mix the right channel audio signal part resulting from the processing of the input audio signal with the right channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a right channel output signal.

[0028] The method thus, allows for implementing a Hybrid Mono-Stereo Uniform Partition Overlap-Save ("HUPOLS") reverberation principle. The HUPOLS principle is an efficient method of convolving an input audio signal - which can be a mono or a stereo audio signal - with a real-life room stereo impulse response of large length. The HUPOLS principle is based on the conventional UPOLS method and the generic principle of performing concurrent stereo convolution of the input audio signal with the early stereo part of the RIR and mono convolution of the input audio signal with the late mono part of the RIR. HUPOLS thereby, significantly reduces the number of FLOPS and the amount of memory needed, without any noticeable degradation of the stereo perception of the stereo output audio signal, by exploiting the different effect and importance that the early and the late reflections of a real-life room have on the reverberation.

ated stereo output audio signal. Moreover, HUPOLS can exploit the structure of a UPOLS building block to economize on the Discrete Fourier Transform operations ("DFT" operations) and Inverse Discrete Fourier Transform operations ("IDFT" operations) and to eliminate the delay needed for modelling the late mono part of the RIR.

[0029] Specifically, the method makes use of the insight of taking advantage of the differing subjective perceptions caused by the early and late parts of a pre-recorded RIR. The method particularly, uses a stereo model for the early-reflections-part of the RIR to reproduce the reverberation caused by the early reflections of the room. This early part determines the spatial impression, the understanding of our position, and the sound source position within the room. For the late-reflections-part of the RIR, the method particularly, uses a mono model to reproduce the reverberation caused by the late reflections of the room. This late part determines the perception of the room size and geometry. Given the fact that the early reflections are much shorter in duration compared to the late reflections, the method achieves a noticeable reduction in the required computing resources, since the processing of stereo audio signals requires twice the resources as the processing of mono audio signals. As indicated above, the input audio signal applied to the stereo model of the early-reflections-part of the RIR can be mono or stereo, whereas the generated output signal is always a stereo signal. If the duration of the early and the late part of the RIR is properly determined, the method allows for reducing the required computing resources at no expense in the quality of the reverberated stereo audio signal.

[0030] Hence, an improved method of processing an input audio signal for generating a stereo output audio signal with specific reverberation characteristics, particularly with respect to the computing power and memory required for the respective digital audio signal processing unit as well as ease of implementation, is given.

[0031] In the following, exemplary embodiments for processing of a mono input audio signal and for processing of a stereo input audio signal will be explained:

For a mono input audio signal, step e) can comprise applying or implementing a or the first signal processing rule for processing, particularly by convolving, the mono input audio signal with the left channel samples of the stereo part of the RIR and for processing, particularly by convolving, the mono input audio signal with the right channel samples of the stereo part of the RIR. Thereby, a processed left audio signal part and a processed right audio signal part representing the reverberation of the mono input audio signal from the first group of samples of the RIR data can be obtained. Further, step e) can comprise applying or implementing a or the second signal processing rule for processing, particularly by convolving, the mono input audio signal with the mono part of the RIR data, thereby obtaining a processed mono audio signal part representing the reverberation of the mono input audio signal from the second group of samples of

the RIR data.

[0032] In this exemplary embodiment, step f) can comprise mixing the processed mono audio signal part with the processed left audio signal part, thereby generating a or the left channel output signal, and mixing the processed mono audio signal part with the processed right audio signal part, thereby generating a or the right channel output signal. Again, the generated left and right channel output signals build the stereo output audio signal having the specific reverberation.

[0033] For a stereo input audio signal, step e) can comprise applying or implementing a or the first signal processing rule for processing, particularly by convolving, the left channel of the stereo input audio signal with the left channel samples of the stereo part of the RIR data and for processing the right channel of the stereo input audio signal with the right channel samples of the stereo part of the RIR data, thereby obtaining a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of the left and right channel of the stereo input audio signal from the left channel samples and from the right channel samples of the stereo part of the RIR. Further, step e) can comprise applying or implementing a or the second signal processing rule for processing, particularly by convolving, the mono version of the stereo input audio signal with the mono part of the RIR data, thereby obtaining a processed mono audio signal part, representing the reverberation of the mono version of the stereo input audio signal from the second group of samples of the RIR.

[0034] In this exemplary embodiment, step f) can comprise mixing the left channel audio signal part resulting from the processing of the input signal with the left channel samples of the stereo part of the RIR data with the mono audio signal part resulting from the processing of the input signal with the mono part of the RIR data, thereby generating a reverberated left channel output signal; and mixing the right channel audio signal part resulting from the processing of the input signal with the right channel samples of the stereo part of the RIR data with the mono audio signal part resulting from the processing of the input signal with the mono part of the RIR data, thereby generating a reverberated right channel output signal. Again, the generated left and right channel output signals build the stereo output audio signal having the specific reverberation.

[0035] In exemplary embodiments, the method can comprise outputting the left channel output signal via a left output audio channel and outputting the right channel output signal via a right output audio channel. Respective left and right output audio channels can be embodied through loudspeakers of an audio system, i.e. particularly a vehicle audio system or a car audio system, i.e. an audio system that is to be installed or is installed in a vehicle or a car.

[0036] In exemplary embodiments in which a stereo input audio signal is processed, the left and right channel

of the stereo input audio signal can be pre-processed by applying a pre-processing rule for converting stereo input audio signal to mono before applying the second signal processing rule. A respective pre-processing rule can be embodied via a hardware- and/or software embodied pre-processing unit which is configured to pre-process the left and right channel of stereo input audio signal for converting stereo input audio signal to mono before applying the second signal processing rule. A respective pre-processing can be beneficial for the (subsequent) application or implementation of the second signal processing rule, e.g. due to reduced computational efforts for carrying out the second signal processing rule.

[0037] A respective pre-processing rule for converting the stereo input audio signal to mono can comprise forming the arithmetic mean between the left channel samples and the right channel samples of the stereo input audio signal. In other words, a respective pre-processing rule for converting the stereo input audio signal to mono can comprise summing the left channel samples with the right channel samples and for each pair of samples that have been added together dividing the result by two.

[0038] The summing can particularly, comprise adding of corresponding blocks of the left and the right channel of the or a respective stereo input audio signal. The summing typically, further comprises or can be followed by dividing the result of the addition by two.

[0039] The method typically, comprises applying a time-delay filter before application of the second signal processing rule for processing the input audio signal with the mono part of the RIR. The time-delay filter can be applied by a hardware- and/or software-embodied time-delay filter unit which is configured to apply a time-delay filter before application of the second signal processing rule.

[0040] The time delay introduced by the time-delay filter is typically, equal to the time duration of the stereo part of the RIR data. As such, the length of the delay filter typically, corresponds to the length of the stereo part of the RIR.

[0041] In exemplary embodiments, the first and second signal processing rule can each comprise at least one filtering operation, particularly at least one convolving operation. Particularly, the first signal processing rule typically, comprises (exactly) two filtering operations and the second signal processing rule typically, comprises (exactly) one filtering operation. As such, the hardware- and/or software-embodied signal processing unit or structure for implementing the first and second signal processing rule can be embodied as filtering units, particularly as convolution units, configured to comprise at least one filtering operation, particularly at least one convolving operation.

[0042] In exemplary embodiments, the determination of the first number of RIR samples representing the stereo part of the RIR and the second number of RIR samples representing the mono part of the RIR can be done iteratively. Generally, the determination of the first number

of RIR samples representing the stereo part of the RIR and the second number of RIR samples representing the mono part of the RIR can be done experimentally, e.g. using a suitable hardware- and software-embodied signal processing structure. This determination can be an iterative process and can require the attention of a user, i.e. particularly an audio engineer.

[0043] As indicated above, a signal processing unit or structure can be used for applying or implementing both the first and the second signal processing rule.

[0044] In exemplary embodiments, if the input audio signal is stereo, the signal processing structure can comprise four hardware- and/or software-embodied signal processing blocks, particularly built as or comprising Discrete Fourier Transform blocks, particularly Fast Discrete Fourier Transform blocks. If the input audio signal is mono, the signal processing structure can comprise three hardware- and/or software-embodied signal processing blocks, particularly built as or comprising Discrete Fourier Transform blocks, particularly Fast Discrete Fourier Transform blocks. The signal processing structure used for implementing the method or the respective steps of the method can thus, have a relatively simple and/or effective configuration.

[0045] Particularly, the signal processing structure can comprise one or more first signal processing blocks for implementing the first signal processing rule, and one or more second signal processing blocks for implementing the second signal processing rule.

[0046] A second aspect of the invention refers to a signal processing device, comprising means, particularly a signal processing structure, for carrying out the method of the first aspect of the invention. Thus, all remarks regarding the method of the first aspect also apply to the signal processing device.

[0047] A third aspect of the invention refers to a computer program product comprising instructions which, when the program is executed by a computer, particularly a DSP unit, cause the computer to carry out the method of the first aspect of the invention. Thus, all remarks regarding the method of the first aspect also apply to the computer program product.

[0048] A fourth aspect of the invention refers to a computer-readable data carrier having stored thereon the computer program product of the third aspect. Thus, all remarks regarding the method of the first aspect also apply to the computer-readable data carrier.

[0049] A fifth aspect of the invention refers to an audio processing apparatus for processing an input audio signal, comprising a signal processing device according to the second aspect. Thus, all remarks regarding the method of the first aspect also apply to the audio processing apparatus.

[0050] A sixth aspect of the invention refers to a vehicle, particularly a car, comprising an audio processing apparatus for processing an input audio signal according to the fifth aspect. Thus, all remarks regarding the method of the first aspect also apply to the vehicle.

[0051] Exemplary embodiments of diverse aspects of the invention are described in context with the following Figures, whereby:

Fig. 1 shows a principle drawing of a digital signal processing structure for implementing a method of processing a stereo input audio signal for generating a stereo output audio signal of a specific reverberation according to an exemplary embodiment of the invention;

Fig. 2 shows a principle drawing of a digital signal processing structure for implementing a method of processing a mono input audio signal for generating a stereo output audio signal of a specific reverberation according to an exemplary embodiment of the invention;

Fig. 3 shows abstract models of the digital signal processing structures of Fig. 1 (see box II) and Fig. 2 (see box III);

Fig. 4 shows a UPOLS-stereo system according to an exemplary embodiment; and

Fig. 5 shows a UPOLS-mono system according to an exemplary embodiment.

[0052] Fig.1 shows a principle drawing of a digital signal processing structure 100 for implementing a method of processing a stereo input audio signal for generating a stereo output audio signal of a specific reverberation according to an exemplary embodiment of the invention.

[0053] The method enables generating a stereo output audio signal having a specific reverberation, i.e. specific reverberation characteristics, such as the specific reverberation characteristics of a specific acoustic environment, e.g. a specific room of a specific building, by processing of an input audio signal.

[0054] The basic steps of the method of processing an input audio signal for generating a stereo output audio signal of a specific reverberation will be specified in the following:

According to a first step of the method, an input audio signal is provided. The input audio signal is a mono input audio signal ("mono signal") or a stereo input audio signal ("stereo signal").

[0055] The first step of the method can be implemented by a hardware- and/or software-embodied input audio signal providing unit which is configured to provide an input audio signal that is mono or stereo. An input audio signal can be a music- and/or speech-signal, i.e. a signal comprising music- and/or speech-content.

[0056] According to a second step of the method, pre-recorded stereo data of a Room-Impulse-Response ("RIR") of a specific acoustic environment, such as a specific building or venue, are provided. The pre-recorded RIR data, respectively comprise a defined number of RIR

samples, i.e. left channel samples and right channel samples. Typically, the pre-recorded RIR data are or comprise stereo data. Particularly, the RIR data comprise an equal number of left channel samples and right channel samples. The RIR data can be obtained through known methods for recording the RIR of acoustic environments. The actual recording of a respective acoustic environment is typically, not a step of the method.

[0057] The second step of the method can be implemented by a hardware- and/or software-embodied RIR data providing unit which is configured to provide pre-recorded RIR data of a specific acoustic environment.

[0058] According to a third step of the method, a first number of RIR samples representing a stereo part of the RIR represented by the RIR data and a second number of RIR samples representing a mono part of the RIR represented by the RIR data is determined. Thereby, the stereo part of the RIR comprises a number of left channel samples for the left output channel and an equal number of right channel samples for the right output channel. The mono part of the RIR represented by the RIR data comprises a number of RIR samples to be used for both the left and the right output channel. The third step of the method thus, comprises the determination of a stereo part of the RIR which is or can be represented by a first number of RIR samples and a mono part of the RIR which is or can be represented by a second number of RIR samples. The stereo part of the RIR is the first part of the RIR and is followed by the mono part of the RIR which is the second part of the RIR. The duration of the stereo part and of the mono part of the RIR typically, add up to the total duration of the RIR.

[0059] The third step of the method can be implemented by a hardware- and/or software-embodied determining unit which is configured to determine a first number of RIR samples representing a stereo part of the RIR and a second number of RIR samples representing a mono part of the RIR, whereby the stereo part of the RIR comprises a number of left channel RIR samples for the left output channel and an equal number of right channel RIR samples for the right output channel. The mono part of the RIR comprises a number of RIR samples for both the left and the right output channel.

[0060] According to a fourth step of the method, the samples of the RIR are subdivided into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono part of the RIR. Hence, the first group of RIR samples typically comprises the first number of RIR samples representing the stereo part of the RIR and the second group of RIR samples typically comprises the second number of RIR samples representing the mono part of the RIR. The first group of RIR samples typically, represents a (distinct) early reflection part of the RIR ("ERP") and the second group of RIR samples typically, represents a (distinct) late reflection part of the RIR ("LRP"). The first group of RIR samples can comprise a period ranging between 1 ms and 150 ms, particularly between

10 ms and 100 ms, of the (initial) duration of the RIR. The second group of RIR samples comprises the remaining duration of the RIR.

[0061] The fourth step of the method can be implemented by a hardware- and/or software-embodied subdividing unit which is configured to subdivide the samples of the RIR into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono part of the RIR.

[0062] The third and fourth step of the method can be combined in one step which comprises both the determining aspect as specified above in context with the third step and the subdividing aspect as specified above in context with the fourth step. Hence, a hardware- and/or software-embodied determining and subdividing unit which is configured to determine a first number of RIR samples representing a stereo part of the RIR and a second number of RIR samples representing a mono part of the RIR, whereby the stereo part of the RIR comprises a number of left channel RIR samples for the left output channel and an equal number of right channel RIR samples for the right output channel and configured to subdivide the samples of the RIR into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono

part of the RIR, can be used when the third and fourth step are combined.

[0063] According to a fifth step of the method, a first signal processing rule for processing, particularly by convolving, the input audio signal with the left channel samples of the stereo part of the RIR and for processing, particularly by convolving, the input audio signal with the right channel samples of the stereo part of the RIR is applied or implemented. Thereby, a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of the input audio signal - which can be a mono input audio signal or a stereo input audio signal - from the first group of samples of the RIR is obtained. Hence, by applying or implementing the first signal processing rule, the input audio signal is processed, i.e. typically convolved, with the left channel samples of the stereo part of the RIR, whereby a processed, i.e. typically convolved, left channel audio signal part is obtained. Likewise, the input audio signal is processed, i.e. typically convolved, with the right channel samples of the stereo part of the RIR, whereby a processed, i.e. typically convolved, right channel audio signal part is obtained. The respective processed left and right channel audio signal parts represent the reverberation, i.e. specifically the artificially generated reverberation, of the input audio signal from the first group of samples of the RIR.

[0064] Further, the fifth step comprises applying or implementing a second signal processing rule for processing, particularly by convolving, the input audio signal with the mono part of the RIR. Thereby, if the input audio signal is a mono input audio signal, a processed mono audio signal part representing the reverberation of the

mono input audio signal from the second group of samples of the RIR is obtained; and, if the input audio signal is a stereo input audio signal, a processed mono audio signal part representing the reverberation of the mono version of the stereo input audio signal from the second group of samples of the RIR is obtained. Hence, by applying or implementing the second signal processing rule, a mono input audio signal is processed so as to obtain a processed mono audio signal part representing the reverberation of the mono input audio signal from the second group of samples of the RIR (the mono part of the RIR), and a stereo input audio signal is processed so as to obtain a processed mono audio signal part representing the reverberation of both the left and right channel of the stereo input audio signal from the second group of samples of the RIR (the mono part of the RIR). The signal resulting from the processing of the input audio signal with the mono part of the RIR, namely with the second group of samples of the RIR, is always a mono signal, regardless if the input audio signal is a mono or a stereo signal. Thus, the mono input audio signal, or the stereo input audio signal after being converted into a mono input audio signal, is processed with the second group of samples of the RIR, namely with the mono part of the RIR, to generate a processed mono audio signal, representing the reverberation of the input audio signal with the second group of samples of the RIR, namely with the mono part of the RIR.

[0065] The fifth step of the method can be implemented by a hardware- and/or software-embodied signal processing structure or unit which is configured to apply a first signal processing rule for processing, particularly by convolving, the input audio signal with the left channel samples of the stereo part of the RIR data and for processing, particularly by convolving, the input audio signal with the right channel samples of the stereo part of the RIR data, thereby obtaining a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of the mono or stereo input audio signal from the first group of samples of the RIR. Further, the hardware- and/or software-embodied signal processing structure or unit is configured to apply a second signal processing rule for processing, particularly by convolving, the input audio signal with the mono part of the RIR data, thereby obtaining a processed mono audio signal part representing the reverberation of the mono or stereo input audio signal from the second group of samples of the RIR.

[0066] According to a sixth step of the method, the left channel audio signal part resulting from the processing of the input audio signal with the left channel samples of the stereo part of the RIR is mixed with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, whereby a left channel output signal is generated. Further, the right channel audio signal part resulting from the processing of the input audio signal with the right channel samples of the stereo part of the RIR is mixed with the audio signal part resulting

from the processing of the input audio signal with the mono part of the RIR, whereby a right channel output signal is generated. Hence, by mixing the left channel audio signal part with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, a left channel output signal is generated; and by mixing the right channel audio signal part with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, a right channel output signal is generated. The generated left and right channel output signals build the stereo output audio signal having the specific reverberation characteristics.

[0067] The sixth step of the method can be implemented by a hardware- and/or software-embodied mixing unit which is configured to mix the left channel audio signal part resulting from the processing of the input audio signal with the left channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a left channel output signal; and to mix the right channel audio signal part resulting from the processing of the input audio signal with the right channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a right channel output signal.

[0068] In the following, exemplary embodiments for processing of a mono input audio signal and for processing of a stereo input audio signal will be explained:

For a mono input audio signal, step e) can comprise applying or implementing a or the first signal processing rule for processing, particularly by convolving, the mono input audio signal with the left channel samples of the stereo part of the RIR and for processing, particularly by convolving, the mono input audio signal with the right channel samples of the stereo part of the RIR. Thereby, a processed left audio signal part and a processed right audio signal part representing the reverberation of the mono input audio signal from the first group of samples of the RIR can be obtained. Further, step e) can comprise applying or implementing a or the second signal processing rule for processing, particularly by convolving, the mono input audio signal with the mono part of the RIR data, thereby obtaining a processed mono audio signal part representing the reverberation of the mono input audio signal from the second group of samples of the RIR data.

[0069] In this embodiment, step f) can comprise mixing the processed mono audio signal part with the processed left audio signal part, thereby generating a or the left channel output signal, and mixing the processed mono audio signal part with the processed right audio signal part, thereby generating a or the right channel output signal. Again, the generated left and right channel output signals build the stereo output audio signal having the specific reverberation.

[0070] For a stereo input audio signal, step e) can com-

prise applying or implementing a or the first signal processing rule for processing, particularly by convolving, the left channel of the stereo input audio signal with the left channel samples of the stereo part of the RIR data and for processing, particularly by convolving, the right channel of the stereo input audio signal with the right channel samples of the stereo part of the RIR data, thereby obtaining a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of the left and right channel of the stereo input audio signal from the left channel samples and from the right channel samples of the stereo part of the RIR. Further, step e) can comprise applying or implementing a or the second signal processing rule for processing, particularly by convolving, the mono version of the stereo input audio signal with the mono part of the RIR data, thereby obtaining a processed mono audio signal part representing the reverberation of the mono version of the stereo input audio signal from the second group of samples of the RIR.

[0071] In this embodiment, step f) can comprise mixing the left channel audio signal part resulting from the processing of the input signal with the left channel samples of the stereo part of the RIR data with the mono audio signal part resulting from the processing of the input signal with the mono part of the RIR data, thereby generating a reverberated left channel output signal; and mixing the right channel audio signal part resulting from the processing of the input signal with the right channel samples of the stereo part of the RIR data with the mono audio signal part resulting from the processing of the input signal with the mono part of the RIR data, thereby generating a reverberated right channel output signal. Again, the generated left and right channel output signals build the stereo output audio signal having the specific reverberation.

[0072] In exemplary embodiments, the method can comprise outputting the left channel output signal via a left output audio channel and outputting the right channel output signal via a right output audio channel. Respective left and right output audio channels can be embodied through loudspeakers of an audio system, i.e. particularly a vehicle audio system or a car audio system, i.e. an audio system that is to be installed or is installed in a vehicle or a car.

[0073] In exemplary embodiments in which a stereo input audio signal is processed, the left and right channel of the stereo input audio signal can be pre-processed by applying a pre-processing rule for converting stereo input audio signal to mono before applying the second signal processing rule. A respective pre-processing rule can be embodied via a hardware- and/or software embodied pre-processing unit which is configured to pre-process the left and right channel of the stereo input audio signal for converting the stereo input audio signal to mono before applying the second signal processing rule. A respective pre-processing can be beneficial for the (subsequent) application or implementation of the second sig-

nal processing rule, e.g. due to reduced computational efforts for carrying out the second signal processing rule.

[0074] A respective pre-processing rule for converting the stereo input audio signal to mono can comprise forming the arithmetic mean between the left channel samples and the right channel samples of the stereo input audio signal. In other words, a respective pre-processing rule for converting the stereo input audio signal to mono can comprise summing the left channel samples with the right channel samples and for each pair of samples that have been added together dividing the result by two.

[0075] The summing can particularly, comprise adding of corresponding blocks of the left and the right channel of the or a respective stereo input audio signal. The summing typically, further comprises or can be followed by dividing the result of the addition by two.

[0076] The method can further comprise applying a time-delay filter before application of the second signal processing rule for processing the input audio signal with the mono part of the RIR. The time-delay filter can be applied by a hardware- and/or software-embodied time-delay filter unit which is configured to apply a time-delay filter before application of the second signal processing rule.

[0077] The time delay introduced by the time-delay filter is typically, equal to the time duration of the stereo part of the RIR data. As such, the length of the delay filter typically, corresponds to the length of the stereo part of the RIR.

[0078] The first and second signal processing rule can each comprise at least one filtering operation, particularly at least one convolving operation. Particularly, the first signal processing rule typically, comprises (exactly) two filtering operations and the second signal processing rule typically, comprises (exactly) one filtering operation. As such, the hardware- and/or software-embodied signal processing unit or structure for implementing the first and second signal processing rule can be embodied as filtering units, particularly as convolution units, configured to comprise at least one filtering operation, particularly at least one convolving operation.

[0079] The determination of the first number of RIR samples representing the stereo part of the RIR and the second number of RIR samples representing the mono part of the RIR can be done iteratively. Generally, the determination of the first number of RIR samples representing the stereo part of the RIR and the second number of RIR samples representing the mono part of the RIR can be done experimentally, e.g. using a suitable hardware- and software-embodied signal processing structure. This determination can be an iterative process and can require the attention of a user, i.e. particularly an audio engineer.

[0080] As is apparent from above, the exemplary embodiments of the method comprise subdividing a pre-recorded RIR in early-stereo blocks (first group of RIR samples) and in late-mono blocks (second group of RIR samples).

[0081] According to exemplary embodiments, if the number of samples L that the RIR sequence h_n , $0 \leq n < L$, contains is not an integer multiple of the block size B , then the minimum required number of trailing zero samples are appended to the sequence h_n to extend it to a length that is a multiple of B . These zero samples are placed after the last sample of the sequence. The length of the sequence increases then to $B[L/B]$, where $[L/B]$ is the smallest integer greater than or equal to (L/B) . The samples of the resulting sequence are then partitioned into K blocks h_k , each having B samples, where h_k is defined as $h_k = [h_{kB+0}, h_{kB+1}, \dots, h_{kB+(B-1)}]_{1 \times B}$, $0 \leq k < K$, $K = L/B$, and L the length of the sequence after the zero-padding (if any). For the smallest possible block size of $B = 1$ sample, no trailing zeros are appended to the RIR. In this case $K = L$. The case $B = 1$ is a marginal case of no practical interest. For a stereo RIR, h_n and h_k refer to any of the left or the right channel. For a stereo RIR the handling described above is applied to the sequence h_n^L for the left channel and to the sequence h_n^R for the right channel of the RIR, resulting the K blocks h_k^L defined as $h_k^L = [h_{kB+0}^L, h_{kB+1}^L, \dots, h_{kB+(B-1)}^L]_{1 \times B}$ for the left channel and the K blocks h_k^R defined as $h_k^R = [h_{kB+0}^R, h_{kB+1}^R, \dots, h_{kB+(B-1)}^R]_{1 \times B}$ for the right channel, where as before, $0 \leq k < K$, $K = L/B$, and L the size of the sequences after the zero-padding (if any).

[0082] Considering that the first S blocks of the RIR comprise the early-stereo blocks and that the next and last M blocks of the RIR comprise the late-mono blocks, where $S + M = K$, one can define the early-stereo blocks of the RIR and the late-mono blocks of the RIR as follows:

- the S early-stereo blocks for the left channel are e_k^L , defined as $e_k^L = h_k^L$, $0 \leq k < S$;
- the S early-stereo blocks for the right channel are e_k^R , defined as $e_k^R = h_k^R$, $0 \leq k < S$;
- the M late-mono blocks for a stereo input audio signal are l_k defined as $l_k = (h_{k+S}^L + h_{k+S}^R) / 4$, $0 \leq k < M$;
- the M late-mono blocks for a mono input audio signal are l_k defined as $l_k = (h_{k+S}^L + h_{k+S}^R) / 2$, $0 \leq k < M$.

[0083] An equivalent way to express the above (note that the indexing of the vectors starts from zero) is the following:

- The first S of the K blocks h_k^L are selected as the S early-stereo blocks for the left channel;
- The first S of the K blocks h_k^R are selected as the S early-stereo blocks for the right channel;
- The remaining last M blocks of h_k^L and the remaining last M blocks of h_k^R are combined together to form the M late-mono blocks. The first late-mono block l_0 is formed by adding the $(S+1)^{th}$ block h_S^L of h_k^L with the $(S+1)^{th}$ block h_S^R of h_k^R and dividing the result by 4. The M^{th} late-mono block l_{M-1} is formed by adding the $(S+M)^{th}$ block h_{S+M-1}^L of h_k^L (last block of h_k^L) with the $(S+M)^{th}$ block h_{S+M-1}^R of h_k^R (last block of h_k^R) and dividing the result by 4. It is similar for all

other late-mono blocks. For a mono input audio signal, factor 4 is replaced by 2.

[0084] In the above, the symbol M stands for "mono" and the symbol S stands for "stereo"; the symbol e stands for "early" and the symbol l stands for "late".

[0085] The partitioning of the stereo RIR into early-stereo blocks and late-mono blocks requires the knowledge of the values of parameters S and M, where $1 \leq S < K$, $1 \leq M < K$ (at least one early-stereo block per channel and at least one late-mono block), and $S+M=K$, since all RIR blocks must be modelled.

[0086] An example of the RIR partitioning into early-stereo blocks and late-mono blocks is shown in box I of Fig. 3. In this Fig., the Early Reflections Part ("ERP") of the RIR comprises the early-stereo blocks and the Late Reflections Part ("LRP") of the RIR comprises the late-mono blocks. For the example shown in box I of Fig. 3 it is $S=5$ (5 early-stereo blocks per channel), $M=15$ (15 late-mono blocks), and $K=20$ (a total of 20 blocks per channel comprise the RIR).

[0087] Reference is now made to Fig. 1 which shows a principle drawing of a digital signal processing structure 100 for implementing a method of processing a stereo input audio signal for generating a stereo output audio signal of a specific reverberation according to an exemplary embodiment of the invention;

[0088] According to a high-level description, the digital signal processing structure 100 is divided to the upper part (what stands above block 24) and to the lower part (what stands below block 24). The lower part can be referred to as the "Mono Subsystem" because the signals flowing through it are monophonic, whereas the upper part can be referred to as the "Stereo Subsystem" because the signals flowing through it are stereophonic.

[0089] The "Stereo Subsystem" appears fully symmetric, with its left part implementing the processing for the left channel of the input audio signal and with its right part implementing the processing for the right channel of the input audio signal. Due to the symmetry, blocks for the left and the right channel having the same role also have the same labelling. Block 24 provides the transition from the "Stereo Subsystem" to the "Mono Subsystem" (note its two input ports and one output port). Blocks 27 merge the two subsystems together and provide the transition from the "Mono Subsystem" back to the "Stereo Subsystem". Blocks 15 represent the input to the structure for the left and right channel and blocks 31 the corresponding output of the structure for the left and right channel. The building blocks of the "Stereo Subsystem" can be two modified UPOLS algorithms, one for the left and one for the right channel. These can be referred to as the UPOLS-left and the UPOLS-right subsystems. The building block of the "Mono Subsystem" is a pruned UPOLS method that shares certain blocks with the UPOLS-left and UPOLS-right subsystems. This can be referred to as the UPOLS-mono subsystem.

[0090] The HUPOLS-stereo reverberator illustrated in

Fig. 1 is based on the UPOLS algorithm. The reverberator processes incoming samples x_n frame-by-frame. Here x_n represents the value of the signal at time $n \geq 0$. The signal is assumed to be zero for $n < 0$ and therefore no processing takes place before time zero. The frame size is B samples, where $B \geq 1$. The k^{th} frame to be processed, where $k \geq 0$ is the frame index and frame 0 is the first frame, is the vector of samples $\mathbf{x}_k^L = [x_{kB+0}^L, x_{kB+1}^L, \dots, x_{kB+(B-1)}^L]_{1 \times B}$ and $\mathbf{x}_k^R = [x_{kB+0}^R, x_{kB+1}^R, \dots, x_{kB+(B-1)}^R]_{1 \times B}$ for the left and the right channel, respectively. Buffers 15 contain these samples. Buffer 15 on the left contains the samples of the vector $\mathbf{x}_k^L$ and buffer 15 on the right contains the samples of the vector $\mathbf{x}_k^R$.

[0091] Since in the "Stereo Subsystem" of the HUPOLS-stereo structure the processing on the left (for the left channel) and on the right (for the right channel) is done in exactly the same way, in what follows we only refer to the left channel (left part of the "Stereo Subsystem") and we omit the superscript designating the channel. The first sample x_{kB} (simplified notation for x_{kB}^L) of the vector $\mathbf{x}_k$ (simplified notation for $\mathbf{x}_k^L$) is located at the first (leftmost) location of buffer 15. This convention is followed for all buffers and vectors in this document, namely, the first element of a vector is placed at the first (leftmost) location of the buffer and the last element of a vector is placed at the last (rightmost) location of the buffer.

[0092] For the current new incoming frame $\mathbf{x}_k$ of buffer 15, the samples of the previous frame $\mathbf{x}_{k-1}$ that were present in buffer 16, are shifted to the left by B samples and replace frame $\mathbf{x}_{k-2}$ that was present in buffer 17. In this way, $\mathbf{x}_{k-2}$ is discarded, buffer 17 is filled with $\mathbf{x}_{k-1}$ and buffer 16 is filled with $\mathbf{x}_k$. All these happen during the k^{th} iteration of the algorithm, that determines the output of the algorithm for the current input frame $\mathbf{x}_k$. For the 0th iteration (first iteration) it is set $\mathbf{x}_{-1} = [0, \dots, 0]_{1 \times B}$, which is a vector of B zeros. This means that for the first iteration B zeros are placed in buffer 17 and the vector $\mathbf{x}_0$ in buffer 16. Transform 18 represents the size 2B Real-to-Complex Discrete Fourier Transform ("2B R-C DFT"), of the time-domain vector $[\mathbf{x}_{k-1} | \mathbf{x}_k]$, where $[\mathbf{x}_{k-1} | \mathbf{x}_k] = [x_{kB-B}^L, \dots, x_{kB+(B-1)}^L]_{1 \times 2B}$. This is simply the row-vector formed by the samples of $\mathbf{x}_{k-1}$ (located in buffer 17) followed by the samples of $\mathbf{x}_k$ (located in buffer 16). The output of transform 18 is denoted as $\mathbf{X}_k = [X_{k2B}, \dots, X_{k2B+(2B-1)}]_{1 \times 2B}$. The first Discrete Fourier Transform ("DFT") coefficient (this is the DC term) is X_{k2B} and the last DFT coefficient is $X_{k2B+(2B-1)}$. The input of transform 28 is the frequency-domain vector $\mathbf{Y}_k = [Y_{k2B}, \dots, Y_{k2B+(2B-1)}]_{1 \times 2B}$ and the output of transform 28 is the time-domain vector $[\mathbf{d}_k | \mathbf{y}_k]$, where $[\mathbf{d}_k | \mathbf{y}_k] = [d_{kB}, \dots, d_{kB+(B-1)}]_{1 \times 2B}$. Transform 28 represents the size 2B Complex-to-Real Inverse Discrete Fourier Transform ("2B C-R IDFT") The elements of vector $\mathbf{d}_k = [d_{kB}, \dots, d_{kB+(B-1)}]_{1 \times B}$ are collected in buffer 29 and are discarded (they are not used). The elements of vector $\mathbf{y}_k = [y_{kB+0}, y_{kB+1}, \dots, y_{kB+(B-1)}]_{1 \times B}$ are collected in buffer 30 and form the output of the algorithm to the input frame $\mathbf{x}_k = [x_{kB+0}, x_{kB+1}, \dots, x_{kB+(B-1)}]_{1 \times B}$.

Hence, at time $n \geq 0$, $y_n = x_{kB+m}$ is the output of the algorithm to the input $x_n = x_{kB+m}$, where $0 \leq m < B$ and $k \geq 0$ is the frame index. This output has an inherent delay of $(B-1)$ samples since a total of B input samples need to be collected to build up the block x_k for the processing to start. Typically, only when an input block is complete can the output to this block be calculated. Hence, the latency of the algorithm is typically, $(B-1)$ samples.

[0093] Buffer 1 contains the samples of the vector e_0 (simplified notation for e_{L_0}), buffer 2 the samples of the vector e_1 (simplified notation for e_{L_1}), and buffer 3 the samples of the vector e_{S-1} (simplified notation for $e_{L_{S-1}}$). Buffer 4 contains $\mathbf{0}_B = [0, \dots, 0]_{1 \times B}$, which is a vector of B zeros. Transform 5 represents the size $2B$ R-C DFT of the vector $[e_0 | \mathbf{0}_B]_{1 \times 2B}$. This is the vector formed by the samples of e_0 followed by the samples of $\mathbf{0}_B$. There are S transforms similar to transform 5, for converting the time-domain vectors $[e_k | \mathbf{0}_B]_{1 \times 2B}$, where $0 \leq k < S$, into the frequency-domain vectors $\mathbf{E}_k = [E_{k2B+0}, E_{k2B+1}, \dots, E_{k2B+(2B-1)}]_{1 \times 2B}$, where $0 \leq k < S$. In this notation, E_{k2B} is the first DFT coefficient (the DC term). Buffer 6 contains the first $(B+1)$ elements of the vector $\mathbf{E}_0$. The last $(B-1)$ elements of $\mathbf{E}_0$ are implied by the complex-conjugate symmetry property of the $2B$ R-C DFT and are therefore discarded. There are S buffers of the same size $(B+1)$ as buffer 6, containing the first $(B+1)$ elements of the vectors $\mathbf{E}_k$. Buffers and transforms not shown in Fig. 1 are implied by the ellipsis 7. Buffer 6 and all the buffers underneath are referred to as the Frequency-Domain Left RIR ("FD-LRIR"). A total of S buffers containing complex data comprise the FD-LRIR. The content of the FD-LRIR is typically, calculated off-line and typically, stays constant throughout the streaming and processing of the data. The static FD-LRIR can be stored in the processor memory. The same applies to the Frequency-Domain Right RIR ("FD-RRIR") that appears on the right-hand side of the HUPOLS-stereo structure illustrated in Fig. 1.

[0094] The meaning of buffers and transforms 8 through 14 is similar to that for the buffers and transforms 1 through 7 explained above. Buffer 8 contains the samples of the vector l_0 , buffer 9 the samples of the vector l_1 , and buffer 10 the samples of the vector l_{M-1} . Buffer 11 contains $\mathbf{0}_B = [0, \dots, 0]_{1 \times B}$. Transform 12 represents the size $2B$ R-C DFT of the vector $[l_0 | \mathbf{0}_B]_{1 \times 2B}$. There are M transforms similar to transform 12, for converting the time-domain vectors $[l_k | \mathbf{0}_B]_{1 \times 2B}$, where $0 \leq k < M$, into the frequency-domain vectors $\mathbf{L}_k = [L_{k2B+0}, L_{k2B+1}, \dots, L_{k2B+(2B-1)}]_{1 \times 2B}$, where $0 \leq k < M$. Buffer 13 contains the first $(B+1)$ elements of vector $\mathbf{L}_0$. There are M buffers of the same size $(B+1)$ as buffer 13, containing the first $(B+1)$ elements of the vectors $\mathbf{L}_k$. Buffers and transforms not shown in Fig. 1 are implied by the ellipsis 14. Buffer 13 and all the buffers underneath are referred to as the Frequency-Domain Mono RIR ("FD-MRIR"). A total of M buffers containing complex data comprise the FD-MRIR. The content of the FD-MRIR is typically, calculated off-line and typically, stays constant throughout the streaming and processing of the data. The static FD-MRIR can

be stored in the processor memory.

[0095] All S buffers under transform 18 are referred to as the Frequency-Domain Left Vector Delay Line ("FD-LVLDL"). Buffers 19 and 20 are the first and the last buffer of FD-LVLDL. Buffers not shown in Figure 1 are implied by the ellipsis 21. All S buffers have the same size $(B+1)$ and are initialised with zeros. For the incoming frame x_k of buffer 15, the output $\mathbf{X}_k$ of transform 18 is calculated. The last $(B-1)$ elements of $\mathbf{X}_k$ are implied by the complex-conjugate symmetry property of the size $2B$ R-C DFT. These are all discarded immediately after the output of transform 18 is calculated. The remaining $(B+1)$ samples of $\mathbf{X}_k$ are shifted into buffer 19 and the previous elements of buffer 19 are shifted into the next buffer, namely into the buffer just below. The same happens for all buffers of FD-LVLDL. Namely, every time that buffer elements are shifted downwards into any of the FD-LVLDL buffers, the elements of the buffer where the elements of the buffer located above go into, are also shifted into the next buffer. Buffer 22 and all the buffers underneath are referred to as the Frequency-Domain Mono Vector Delay Line ("FD-MVLDL"). The M buffers of the FD-MVLDL have all the same size $(B+1)$ and are initialised with zeros. Buffers not shown in Fig. 1 are implied by the ellipsis 23. Buffer 22 is the first buffer of the FD-MVLDL. This buffer is updated with the sum of the complex vectors contained in buffers 20 of the FD-LVLDL and the FD-RVLDL (the Frequency-Domain Right Vector Delay Line), just before the elements of buffers 20 are updated. Block 24 is responsible for adding these two complex vectors. Apart from the way that its first buffer is fed, FD-MVLDL works just like FD-LVLDL and FD-RVLDL.

[0096] Given that all $(2S+M)$ buffers of FD-LVLDL, FD-RVLDL, and FD-MVLDL are initialized with zeros, and that the buffer elements move from one buffer to the next buffer (the buffer below) in the way described, the calculation of the output frame $\mathbf{y}^L_0$ and $\mathbf{y}^R_0$ for the input frame $\mathbf{x}^L_0$ and $\mathbf{x}^R_0$ is done using the initial zero values in all M buffers of FD-MVLDL, the initial zero values in all $(2S-2)$ buffers under buffers 19 of FD-LVLDL and FD-RVLDL, and the non-zero values in buffers 19 resulting from the transforms 18. The first input frame for which the initial zero values of FD-LVLDL and FD-RVLDL are completely removed is $\mathbf{x}^L_{S-1}$ and $\mathbf{x}^R_{S-1}$. In a similar way, the first input frame for which the initial zero values of FD-MVLDL are completely removed is $\mathbf{x}^L_{S+M-1}$ and $\mathbf{x}^R_{S+M-1}$. For the calculation of the output frame $\mathbf{y}^L_S$ and $\mathbf{y}^R_S$ for the input frame $\mathbf{x}^L_S$ and $\mathbf{x}^R_S$ all buffers of FD-MVLDL, except from buffer 22, contain the initial zero values.

[0097] The complex multiplier 25 forms the complex vector $[E_0 X_{k2B}, E_1 X_{k2B+1}, \dots, E_B X_{k2B+B}]_{1 \times (B+1)}$. Similarly, each of the multipliers under multiplier 25 forms in a similar way the element-by-element complex product between the contents of its corresponding FD-LVLDL buffer and FD-LRIR buffer. There are S such vector products, each of size $(B+1)$ elements, for each of the S multipliers. The resulting S complex vectors (the vector products) are fed to the upper S input ports of the accumulator block

27. In a similar way, the complex multiplier 26 forms the element-by-element complex product between the contents of its corresponding FD-MVDL buffer and FD-MRIR buffer. There are M such vector products, each of size $(B+1)$ elements, for each of the M multipliers. The resulting M complex vectors (the vector products) are fed to the lower M input ports of the accumulator blocks 27.

[0098] The accumulator block 27 adds the $(S+M) = K$ complex vectors, each of size $(B+1)$ elements, which are provided at its upper S and its lower M input ports, to generate a single complex vector of the same size. This is the vector $[Y_{k2B+0}, Y_{k2B+1}, \dots, Y_{k2B+B}]_{1 \times (B+1)}$. This vector is extended from the size $(B+1)$ to the size $2B$ to yield the complex vector $[Y_{k2B+0}, Y_{k2B+1}, \dots, Y_{k2B+B}, Y_{k2B+B+1}, \dots, Y_{k2B+B+(B-1)}]_{1 \times 2B}$. This extension corresponds to the removal of the last $(B-1)$ elements from the results of the transforms 18. The new $(B-1)$ elements needed for the extension are implied by the complex-conjugate symmetry property of the size $2B$ R-C DFT and are calculated as $Y_{k2B+B+k} = Y_{k2B+B-k}^*$, where the asterisk denotes conjugation and $1 \leq k \leq (B-1)$. The extended vector $[Y_{k2B+0}, Y_{k2B+1}, \dots, Y_{k2B+(2B-1)}]_{1 \times 2B}$ becomes the input of transform 28.

[0099] The complete sequence of the events for the calculation of the output block y_k for the current input block x_k is the following:

- (a) The vector $[x_{k-1} | x_k]$ is transformed according to block 18.
- (b) FD-MVDL is updated from block 24. After this update block 22 contains the sum of the vectors contained in blocks 20 of the left and the right side.
- (c) FD-LVDL is updated from the result of transform 18.
- (d) The vector products for the multipliers 25 and 26 and for all other multipliers under multipliers 25 and 26 are calculated and provided as input to block 27.
- (e) The output of block 27 is calculated and then transformed according to block 28.
- (f) The output block for the left channel is the second half of the transformation result. The processing sequence (a)-(f) is applied for both the left and the right channel concurrently, meaning that every step for the left channel is immediately followed by the corresponding step for the right channel. The processing done for the Mono Subsystem is an exception to this rule, since this subsystem is common for both the left and the right channel.

[0100] Fig. 2 shows another digital signal processing structure 100 according to an exemplary embodiment. The digital signal processing structure can be deemed a HUPOLS-mono reverberator useable or used for a mono input audio signal. Here there is only one input channel and therefore, both FD-LVDL and FD-RVDL of the digital signal processing structure 100 of Fig. 1 are replaced by FD-CVDL ("Frequency-Domain Common Vector Delay Line"). The S buffers comprising the FD-CVDL are buffers

19, 20, and the buffers below buffer 19 and above buffer 20. The first buffer of FD-MVDL is updated directly from the last buffer of FD-CVDL. The vector data of FD-CVDL are used for both the left and the right channel of the ERP of the RIR. Apart from these differences, the HUPOLS-stereo and HUPOLS-mono structures of Fig. 1 and Fig. 2 work in exactly the same way.

[0101] The HUPOLS-stereo reverberator of Fig. 1 achieves the complexity reduction by modelling the last M of the K RIR blocks with the UPOLS-mono subsystem (Mono Subsystem). The UPOLS-left and UPOLS-right subsystems (on the left and on the right of the Stereo Subsystem) model the first S blocks of the left and right channel of RIR. It is $S + M = K$. For $M = 0$, the UPOLS-mono subsystem, the block 24, and the lower M ports of the two blocks 27 in Fig. 1 vanish. The digital signal processing structure 100 of Fig. 1 turns then into the UPOLS-stereo system, that independently processes the left and the right channel of the input audio signal with the left and the right channel of RIR. This digital signal processing structure 100 which can also be denoted UPOLS-stereo system is illustrated in the exemplary embodiment of Fig. 4.

[0102] For $L \gg B \gg 1$, namely, for RIR models of large rooms and for block sizes of acceptable latency but large enough to allow for efficient DFT and IDFT implementations, the resources required for the transforms 18 and 28 in Fig. 1 are negligible compared to the resources needed for the spectral convolutions implemented by blocks 27 and all the multipliers starting from multipliers 25 and 26. Under these conditions, HUPOLS-stereo uses $((2S+M) / (2K)) \times 100\%$ of the resources (FLOPIS and memory) required by the UPOLS-stereo system. This is a number between 50% and 100%. For the marginal case $M=0$ and $S=K$, HUPOLS-stereo uses 100% of the resources required by the UPOLS-stereo system, since the two methods become then identical. For the marginal case $M=K$ and $S=0$ on the other hand, HUPOLS-stereo uses 50% of the resources required by the UPOLS-stereo system. This setting corresponds to a monophonic configuration.

[0103] In a similar way and for $M=0$, in the digital signal processing structure 100 of Fig. 2 the UPOLS-mono subsystem and the lower M ports of the two blocks 27 vanish. The structure of Fig. 2 turns then into the digital signal processing structure 100 illustrated in Fig. 5 which can be deemed a UPOLS-mono system. This system processes the mono input audio signal with the left and the right channel of RIR. HUPOLS-mono uses $((2S+M) / (2K)) \times 100\%$ of the FLOPIS and $((3S+2M) / (3K)) \times 100\%$ of the memory required by the UPOLS-mono system. The last figure ranges from 66.6% to 100%.

[0104] Reference is now made to Fig. 3 which shows abstract models for the digital signal processing structures of Fig. 1 and Fig. 2 in accordance with an exemplary embodiment. In Fig. 3, box II shows the abstract model for the structure of Fig. 1, and box III shows the abstract model for the structure of Fig. 2. Box I of Fig. 3 shows

the partitioning of the RIR into early-stereo and late-mono blocks, that was assumed for box II and box III of this figure.

[0105] For the assumed scenario of box I of Fig. 3, the RIR samples span $K=20$ blocks per channel, each of size B samples. The ERP has $S=5$ blocks per channel and the LRP has $M=15$ blocks. The stereo digital signal processing structure 100 of Fig. 1 and the mono digital signal processing structure 100 of Fig. 2 uses for this configuration 62,5% of the resources required by the UPOLS-stereo system of Fig. 4 or the UPOLS-mono system of Fig. 5. In box I of Fig. 3, blocks starting with block 1 represent the early-stereo blocks e^L_k , $0 \leq k < 5$, for the left channel. Blocks starting with block 2 represent the early-stereo blocks e^R_k , $0 \leq k < 5$, for the right channel. Blocks starting with block 3 represent the late-mono blocks l_k , $0 \leq k < 15$. These blocks were defined above. Blocks starting with block 31 represent the blocks h^L_{k+S} used to define the blocks l_k (see above), where $0 \leq k < 15$. Blocks starting with block 32 represent the blocks h^R_{k+S} used to define the blocks l_k (see above), where $0 \leq k < 15$.

[0106] In box II and box III of Fig. 3, block 4 represents any possible way of convolving the block's input signal with the left channel samples of ERP. It is accordingly for block 5 for the right channel samples of ERP and for block 6 for the mono samples of LRP. Block 7 represents a delay of $SB = 5B$ samples that is needed to time-align the samples of LRP to those of ERP. Blocks 8 and 9 correspond to blocks 15 for the left and the right channel in the HUPOLS-stereo structure of Fig. 1. Block 89 corresponds to block 15 in the HUPOLS-mono structure of Fig. 2. Blocks 10 and 11 correspond to blocks 31 in Fig. 1 and Fig. 2. Adder 12 converts the stereo input audio signal to mono by adding the samples of the left and right channel. The division by two required for this conversion is incorporated into the definition of the late-mono blocks and is therefore omitted from the flow-graph of Fig. 3. Adders 13 and 14 mix the mono output signal of block 6 to the left channel signal (output of block 4) and to the right channel signal (output of block 5), to yield the left channel output (block 10) and the right channel output (block 11) of the model. For stereo configurations, hereinafter denoted as "HUPOLS-s", (box II of Fig.3), the outputs of blocks 4/5 represent the reverberation of the left/right channel of the input audio signal from the early-stereo samples of the left/right channel of RIR. The output of block 6 represents the reverberation of the mono version of the input stereo signal from the late-mono samples of RIR. For mono configurations, hereinafter denoted as "HUPOLS-m", (box III of Fig.3), the left and right channels of the input audio signal are the same (mono input audio signal). Apart from this, the description is the same as for the stereo configurations HUPOLS-s, (box II of Fig.3).

[0107] The HUPOLS-s structures illustrated in Fig. 1 and Fig. 3 (box II) are equivalent, in that they produce the same outputs for the same inputs. Due to this equivalence, reference is made to the digital signal processing

structure 100 of Fig. 3 (box II) as the abstract model of the HUPOLS-s stereo configuration 100 of Fig. 1. The adder 12 of the abstract model corresponds to the vector summation block 24 of HUPOLS-s of Fig. 1. The SB samples delay of the abstract model is implemented in HUPOLS-s of Fig. 1 by the FD-LVDL and FD-RVDL, by the mechanism of buffers 16 and 17, and by the transforms 18. Each buffer of the FD-LVDL corresponds to a delay of B samples and in this way the cascade of the S buffers yields the SB samples delay. The adder 12 of the abstract model is implemented in the stereo configuration of HUPOLS-s of Fig. 1 in the frequency-domain, due to the transform 18 of HUPOLS-s, and is done after the SB samples delay, since the HUPOLS-s block 24 is placed after the FD-LVDL and FD-RVDL. The adders 13 and 14 of the abstract model are implemented with the accumulation blocks 27 of the stereo configuration HUPOLS-s of Fig. 1 for the left and right channel, respectively. Specifically, the effect of adder 13 is achieved by adding the sum of the lower M input vectors to the sum of the upper S input vectors with the left block 27. It is accordingly for adder 14. The adders 13 and 14 of the abstract model are implemented in the stereo configuration HUPOLS-s of Fig. 1 in the frequency-domain, since the left and right blocks 27 stand in-between the transforms 18 and 28. The blocks 4 and 5 of the abstract model correspond to the UPOLS-left and the UPOLS-right subsystems of HUPOLS-s of Fig. 1. Finally, block 6 of the abstract model corresponds to the UPOLS-mono subsystem of Fig. 1. The HUPOLS-m structures illustrated in Fig. 2 and Fig. 3 (box III) are equivalent, in that they produce the same outputs for the same inputs. This equivalence can be explained in a same way as before. Due to this equivalence, reference is made to the digital signal processing structure 100 of Fig. 3 (box III) as the abstract model of the HUPOLS-m stereo configuration 100 of Fig. 2.

[0108] Due to the way that the UPOLS-left, UPOLS-right, and UPOLS-mono subsystems of Fig. 1 interface with each other, HUPOLS-s of Fig. 1 does not require extra memory to implement the delay filter 7 of the abstract model. Moreover, it economizes the DFT and IDFT operations that would normally be needed for implementing block 6 of the abstract model with the stand-alone UPOLS method, by taking advantage of the linearity property of the DFT and IDFT operations. It is similarly for HUPOLS-m of Fig. 2.

[0109] For the HUPOLS-s structure of Fig. 1 the values of parameters S and M , where $1 \leq S < K$, $1 \leq M < K$ and $S + M = K$, must be known. Values of M close to K yield an efficient HUPOLS-s structure, as most of the RIR blocks are then modelled with the UPOLS-mono subsystem but may compromise the stereo quality of the reverberated signal. For RIRs of the same length (with equal number of blocks K), the values of S and M that yield the best trade-off between efficiency and stereo signal quality will generally be different. These values can be found experimentally using the stereo configuration illustrated in Fig. 4. The digital signal processing structure 100 of

Fig. 4 also results from the HUPOLS-s structure of Fig. 1 for $S = K$ and $M = 0$. For this choice of the parameters, block 24, the UPOLS-mono subsystem, and the lower ports of blocks 27 vanish, yielding the UPOLS-stereo system of Fig. 4.

[0110] To evaluate the quality of the output audio signal of the HUPOLS-s structure of Fig. 1 for a certain value of S , where $1 \leq S < K$, the UPOLS-stereo system of Fig. 4 is used. This is operated similarly to the HUPOLS-s structure, but with the following two differences: Immediately after the update of the K buffers of FD-LVDL and FD-RVDL, the two complex vectors $\mathbf{C}^L$ and $\mathbf{C}^R$ contained in buffers $(S+1)$ of FD-LVDL and LFD-RVDL, are both replaced by the complex vector $(\mathbf{C}^L + \mathbf{C}^R)/2$, namely, by their arithmetic mean. As an example, for $S=1$, the vectors contained in buffers 2 (the second buffers) of FD-LVDL and LFD-RVDL are both replaced by their arithmetic mean.

[0111] Immediately after the outputs $\mathbf{V}_k^L$ and $\mathbf{V}_k^R$, where $1 \leq k \leq K$, of the K multipliers for the left and right channel (multipliers 25 and multipliers underneath) are calculated, these outputs for $S+1 \leq k \leq K$ are replaced by their arithmetic mean $(\mathbf{V}_k^L + \mathbf{V}_k^R)/2$, for both the left and the right channel. As an example, for $S=1$, $\mathbf{V}_2^L$ and $\mathbf{V}_2^R$ are replaced by their arithmetic mean. The same is done for $\mathbf{V}_3^L$ and $\mathbf{V}_3^R$, for $\mathbf{V}_4^L$ and $\mathbf{V}_4^R$, etc., and finally for $\mathbf{V}_K^L$ and $\mathbf{V}_K^R$.

[0112] If the above handling is done, the output of the UPOLS-stereo system of Fig. 4 and the output of the HUPOLS-s structure of Fig. 1 configured for the same values of S and M , become identical, provided that the values of S and M remain constant throughout the simulation. In practice, S can vary during the simulation, e.g. can slowly decrease starting from large values, to determine when a deterioration of the stereo signal quality starts being noticeable. Having found the values of S and M for a given RIR, the partitioning of RIR to ERP and LRP can be done as described above and then the HUPOLS-s reverberator of Fig. 1 can be setup as described above.

[0113] For the HUPOLS-m structure of Fig. 2 the values of S and M can be found in a similar way using the UPOLS-mono system of **Fehler! Verweisquelle konnte nicht gefunden werden..** The only difference is that the first step of the handling described before is skipped, since there is only one vector delay line (buffers 19, 20, and all buffers in between), in contrast to the separate vector delay lines for the left and the right channel of Fig. 4 (buffers 19, 20, and all buffers in between), due to the input audio signal being mono.

[0114] The digital signal processing structure 100 of Fig. 1 and Fig. 2 can form part of a hardware- and/or software-embodied digital signal processing device, comprising means, particularly a respective digital signal processing structure 100, for carrying out the method as described in context with the above embodiments.

[0115] A respective digital signal processing device can comprise a computer program product comprising

instructions which, when the program is executed by a computer, cause the computer to carry out the method as described in context with the above embodiments.

[0116] A respective computer program product can be stored on a computer-readable data carrier.

[0117] A respective digital signal processing device can form part of an audio signal processing apparatus for processing an input audio signal.

[0118] A respective audio processing apparatus can be installed in a vehicle, particularly a car.

Claims

1. A method of processing an input audio signal for generating a stereo output audio signal with a specific reverberation, the method comprising the following steps:

- a) providing an input audio signal, the input audio signal being a mono input audio signal or a stereo input audio signal;
- b) providing pre-recorded stereo Room-Impulse-Response ("RIR") data of a specific acoustic environment, the RIR data comprising a defined number of RIR samples, the RIR data comprising an equal number of left channel samples and right channel samples;
- c) determining a first number of RIR samples representing a stereo part of the RIR data and a second number of RIR samples representing a mono part of the RIR data, whereby the stereo part of the RIR data comprises a number of left channel samples for the left output channel and an equal number of right channel samples for the right output channel, and whereby the mono part of the RIR comprises a number of samples to be used for both the left and the right output channel;
- d) subdividing the samples of the RIR into a first group of RIR samples representing the stereo part of the RIR and into a second group of RIR samples representing the mono part of the RIR, whereby the duration that corresponds to the stereo part of the RIR and the duration that corresponds to the mono part of the RIR add up to the total duration of the RIR;
- e) applying a first signal processing rule for processing, particularly by convolving, the input audio signal with the left channel samples of the stereo part of the RIR data and for processing, particularly by convolving, the input audio signal with the right channel samples of the stereo part of the RIR data, thereby obtaining a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of input audio signal from the first group of samples of the RIR, and

- applying a second signal processing rule for processing, particularly by convolving, the mono input audio signal, or the mono version of the stereo input audio signal, with the mono part of the RIR data, thereby obtaining a processed mono audio signal part representing the reverberation of the input audio signal from the second group of samples of the RIR; and
- f) mixing the left channel audio signal part resulting from the processing of the input audio signal with the left channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a left channel output signal; and mixing the right channel audio signal part resulting from the processing of the input audio signal with the right channel samples of the stereo part of the RIR data with the audio signal part resulting from the processing of the input audio signal with the mono part of the RIR data, thereby generating a right channel output signal.
2. The method according to Claim 1, **wherein**, for a mono input audio signal, step e) comprises applying a first signal processing rule for processing, particularly by convolving, the mono input audio signal with the left channel samples of the stereo part of the RIR data and for processing, particularly by convolving, the mono input audio signal with the right channel samples of the stereo part of the RIR data, thereby obtaining a processed left audio signal part and a processed right audio signal part representing the reverberation of the mono input audio signal from the first group of samples of the RIR data, and applying a second signal processing rule for processing, particularly by convolving, the mono input audio signal with the mono part of the RIR data, thereby obtaining a processed mono audio signal part representing the reverberation of the mono input audio signal from the second group of samples of the RIR data; and step f) comprises mixing the processed mono audio signal part with the processed left audio signal part, thereby generating a left channel output signal, and mixing the processed mono audio signal part with the processed right audio signal part, thereby generating a right channel output signal.
3. The method according to Claim 1, **wherein**, for a stereo input audio signal, step e) comprises applying a first signal processing rule for processing, particularly by convolving, the left channel of the stereo input audio signal with the left channel samples of the stereo part of the RIR and for processing, particularly by convolving, the right channel of the stereo input audio signal with the right channel samples of the stereo part of the RIR, thereby obtaining a processed left channel audio signal part and a processed right channel audio signal part representing the reverberation of the left and right channel of the stereo input audio signal from the left channel samples and from the right channel samples of the stereo part of the RIR, and applying a second signal processing rule for processing, particularly by convolving, the mono version of the stereo input audio signal with the mono part of the RIR, thereby obtaining a processed mono audio signal part, representing the reverberation of the mono version of the stereo input audio signal from the second group of samples of the RIR; and step f) comprises mixing the left channel audio signal part resulting from the processing of the left channel of the input audio signal with the left channel samples of the stereo part of the RIR data, with the mono audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, thereby generating a reverberated left channel output signal, and, mixing the right channel audio signal part resulting from the processing of the right channel of the input audio signal with the right channel samples of the stereo part of the RIR data, with the mono audio signal part resulting from the processing of the input audio signal with the mono part of the RIR, thereby generating a reverberated right channel output signal.
4. The method according to any of the preceding Claims, **further comprising** outputting the reverberated left channel output signal via a left output audio channel and outputting the reverberated right channel output signal via a right output audio channel.
5. The method according to any of the preceding Claims, **wherein** for a stereo input audio signal, the left and right channel of the stereo input audio signal is pre-processed by applying a pre-processing rule for converting the stereo input audio signal to mono before applying the second signal processing rule.
6. The method according to Claim 5, **wherein** the pre-processing rule for converting the stereo input audio signal to mono comprises forming the arithmetic mean between the left channel samples and the right channel samples of the stereo input audio signal, whereby each left channel sample is added with its corresponding right channel sample and the result of the addition is divided by two.
7. The method according to any of the preceding Claims, **further comprising** applying a time-delay filter before application of the respective second signal processing rule.
8. The method according to Claim 7, **wherein** the time delay introduced by the time-delay filter is equal to

the time duration of the stereo part of the RIR data.

9. The method according to any of the preceding Claims, **wherein** the first and second signal processing rule comprises a filtering operation, particularly a convolving operation.
10. The method according to any of the preceding Claims, **wherein** the determination of the first number of RIR samples representing the stereo part of the RIR and the determination of the second number of RIR samples representing the mono part of the RIR is done iteratively.
11. The method according to any of the preceding Claims, **wherein** the first group of RIR samples represents a distinct early reflections part of the RIR and the second group of RIR samples represents a distinct late reflections part of the RIR.
12. The method according to any of the preceding Claims, **wherein** the first group of RIR samples comprises a period ranging between 1 ms and 150 ms, particularly between 10 ms and 100 ms, of the initial duration of the RIR data.
13. The method according to any of the preceding Claims, **wherein** a signal processing structure (100) is used for implementing both the first signal processing rule and the second signal processing rule.
14. The method according to Claim 14, **wherein** the signal processing structure (100) comprises at least three signal processing blocks, particularly built as or comprising discrete time Fourier transformation blocks.
15. The method according to Claim 13 or 14, **wherein** the signal processing structure (100) comprises one or more first signal processing blocks, particularly a set of first signal processing blocks, for implementing the first signal processing rule, and one or more second signal processing blocks, particularly a set of second signal processing blocks, for implementing the second signal processing rule.
16. A signal processing device, comprising means, particularly a signal processing structure (100), for carrying out the method of any of the preceding Claims.
17. A computer program product comprising instructions which, when the program is executed by a computer, cause the computer to carry out the method of any of Claims 1 - 15.
18. A computer-readable data carrier having stored thereon the computer program product of claim 17.
19. An audio processing apparatus for processing an input audio signal, comprising a signal processing device according to Claim 16.
20. A vehicle comprising an audio processing apparatus according to Claim 19.

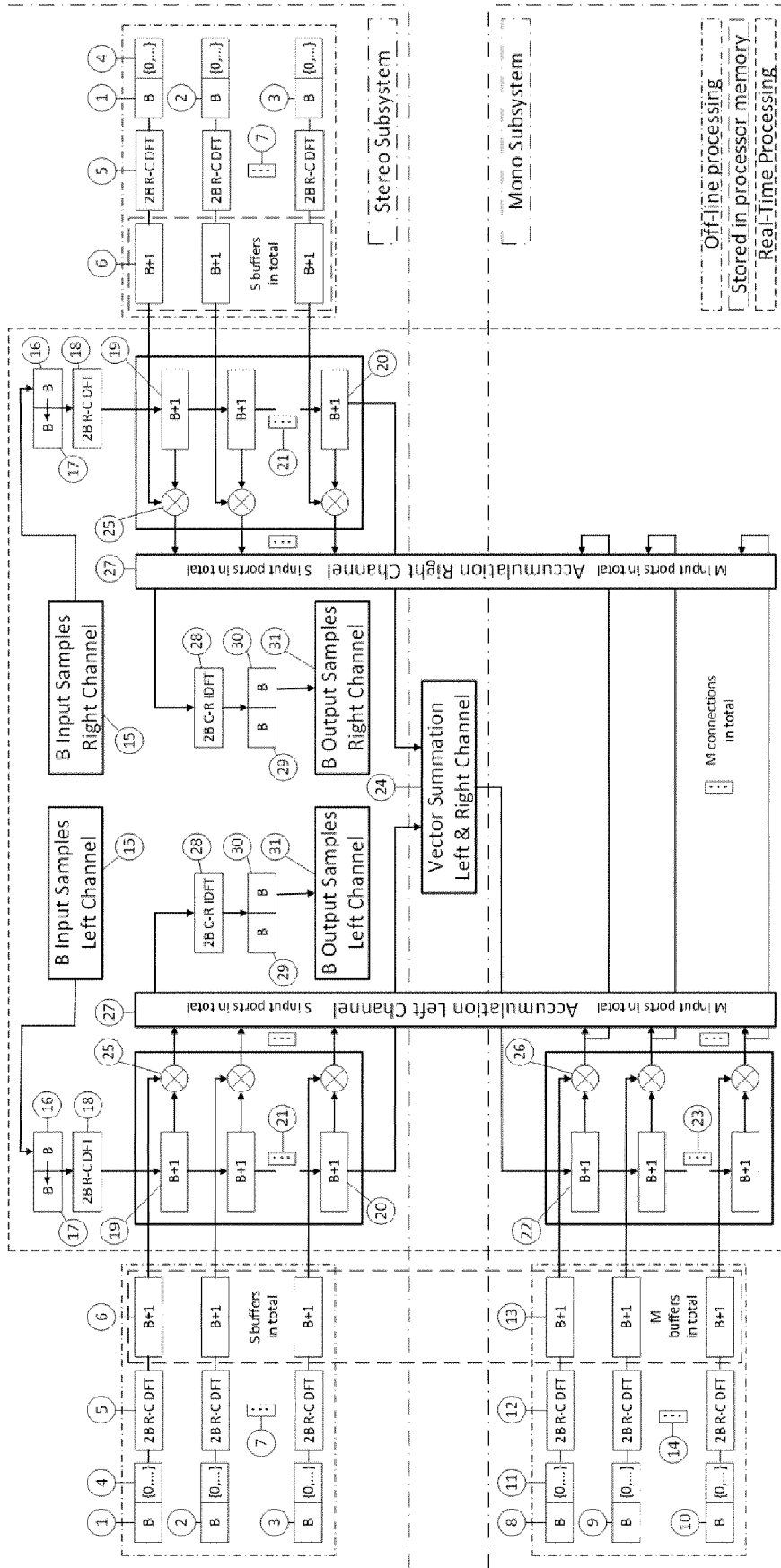


Fig. 1

100

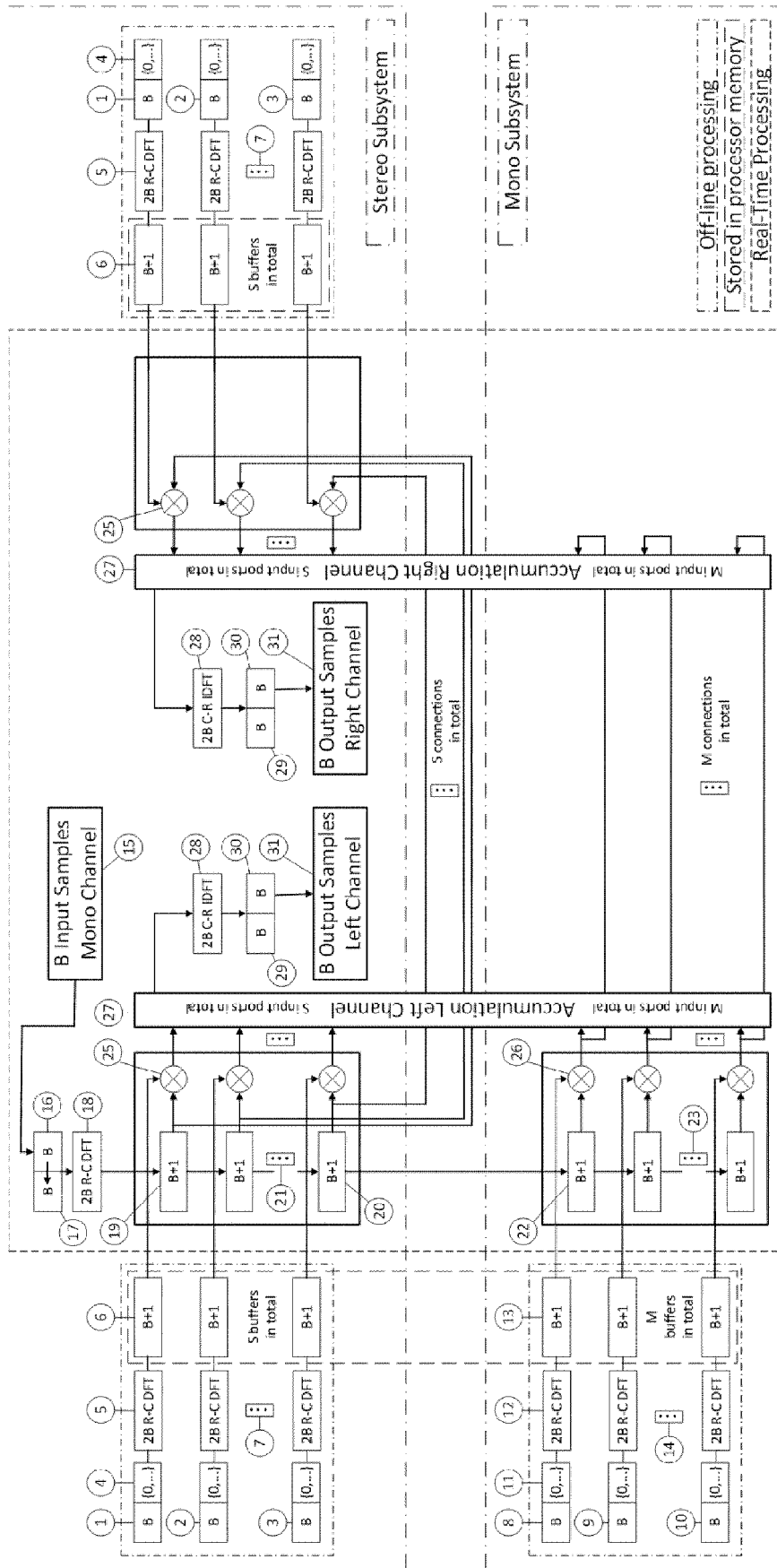


Fig. 2

100

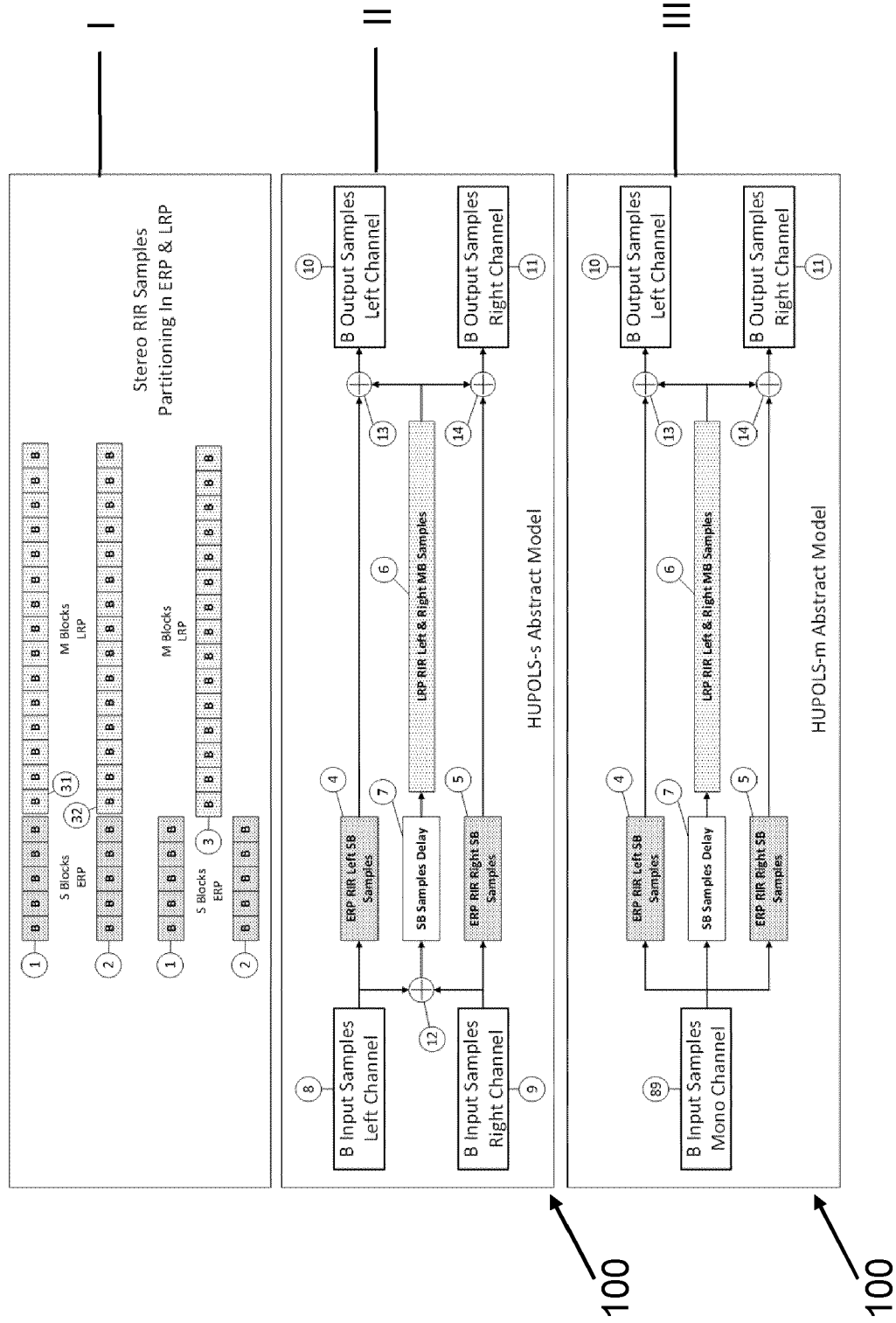


Fig. 3

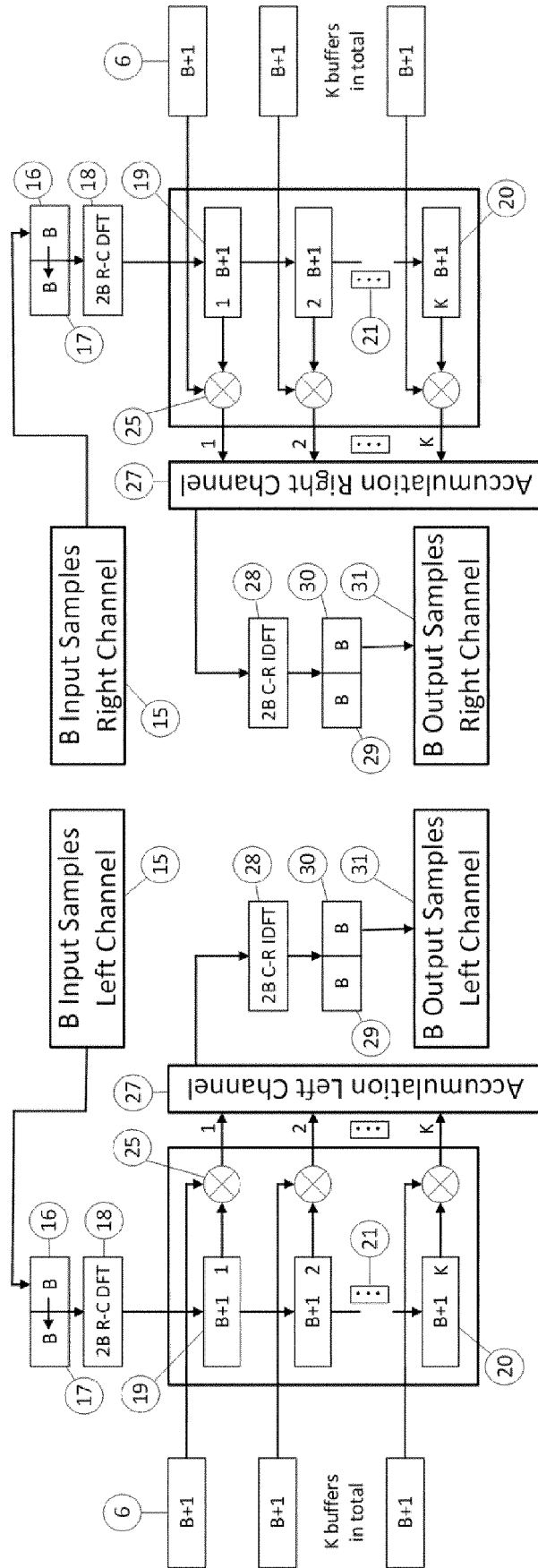


Fig. 4

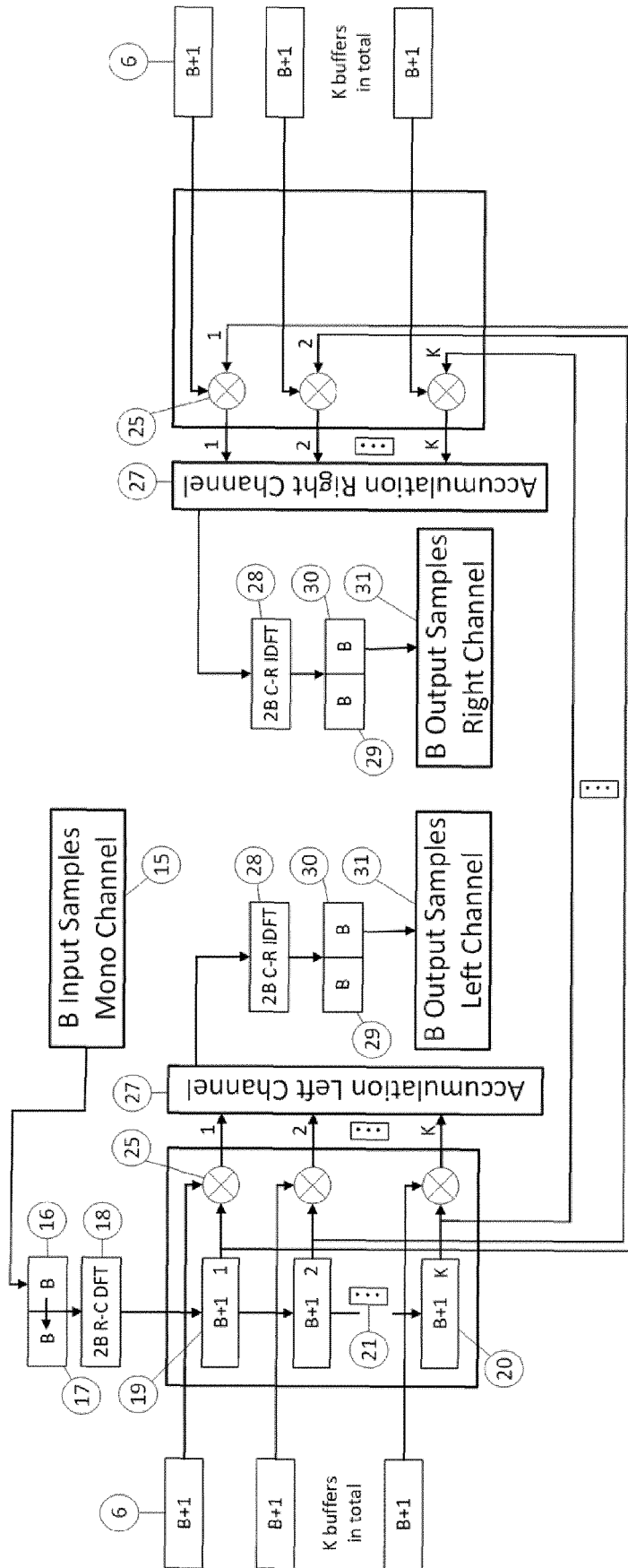


Fig. 5

100



EUROPEAN SEARCH REPORT

Application Number
EP 20 21 0629

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	EP 3 402 222 A1 (DOLBY LABORATORIES LICENSING CORP [US]) 14 November 2018 (2018-11-14) * paragraphs [0001] - [0015], [0044] - [0045]; figures 2,3 *	1-20	INV. H04S1/00 H04S5/00 H04S7/00
X	US 2015/350801 A1 (KOPPENS JEROEN GERARDUS HENRICUS [NL] ET AL) 3 December 2015 (2015-12-03) * figures 10-14 *	1-20	
A	Frank Wefers: "Partitioned convolution algorithms for real-time auralization", 13 May 2015 (2015-05-13), XP055566796, Retrieved from the Internet: URL:http://publications.rwth-aachen.de/record/466561/files/466561.pdf?subformat=pdfa [retrieved on 2019-03-11] * the whole document *	1-20	
			TECHNICAL FIELDS SEARCHED (IPC)
			H04S
The present search report has been drawn up for all claims			
Place of search The Hague		Date of completion of the search 26 April 2021	Examiner Radomirescu, B-M
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document			

 1
EPO FORM 1503 03/82 (P04C01)

**ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.**

EP 20 21 0629

5 This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

26-04-2021

10

Patent document
cited in search report

Publication
date

Patent family
member(s)

Publication
date

15

EP 3402222 A1 14-11-2018

AU 2014374182 A1 30-06-2016

AU 2018203746 A1 21-06-2018

AU 2020203222 A1 04-06-2020

CA 2935339 A1 09-07-2015

CA 3043057 A1 09-07-2015

CN 104768121 A 08-07-2015

EP 3402222 A1 14-11-2018

EP 3806499 A1 14-04-2021

JP 6215478 B2 18-10-2017

JP 2017507525 A 16-03-2017

KR 20160095042 A 10-08-2016

KR 20180071395 A 27-06-2018

MX 352134 B 10-11-2017

RU 2637990 C1 08-12-2017

25

US 2015350801 A1 03-12-2015

BR 112015016978 A2 11-07-2017

CN 104919820 A 16-09-2015

EP 2946572 A1 25-11-2015

JP 6433918 B2 05-12-2018

JP 2016507986 A 10-03-2016

MX 346825 B 03-04-2017

RU 2015134388 A 22-02-2017

US 2015350801 A1 03-12-2015

WO 2014111829 A1 24-07-2014

35

40

45

50

55

ORM P0459