



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
**30.07.2014 Bulletin 2014/31**

(51) Int Cl.:  
**G10L 21/0208** <sup>(2013.01)</sup>

(21) Application number: **13153062.8**

(22) Date of filing: **29.01.2013**

(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**

(71) Applicant: **QNX Software Systems Limited**  
**Kanata, Ontario K2K 0B3 (CA)**

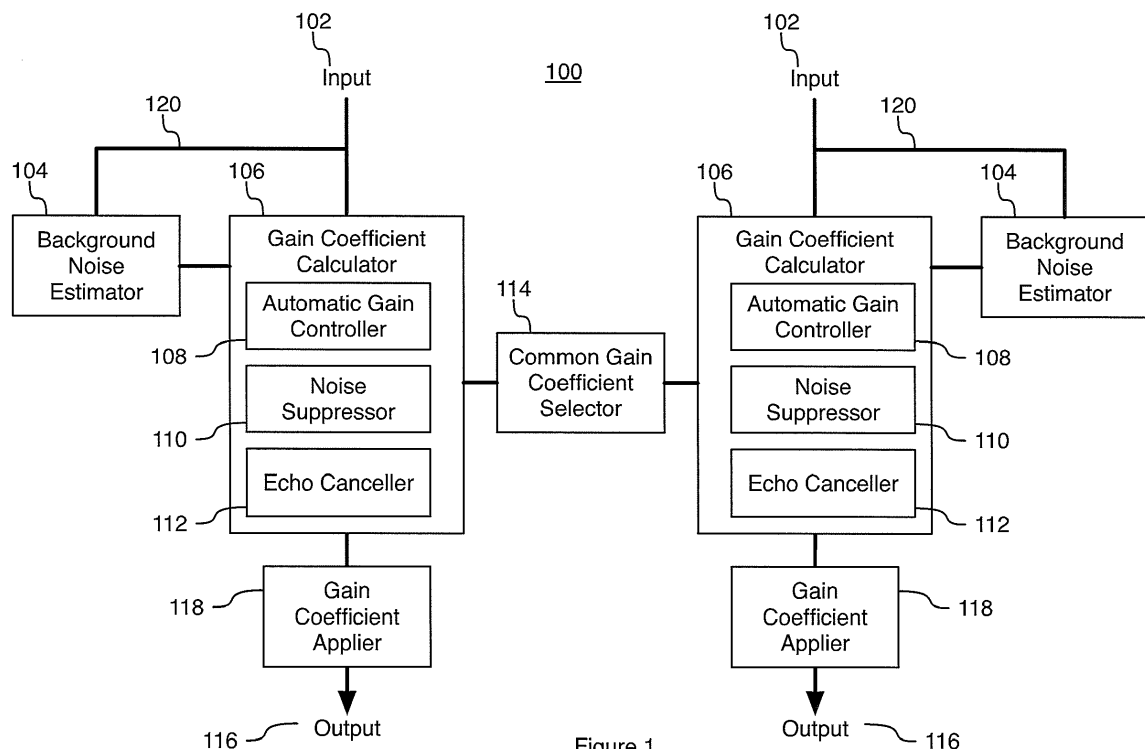
(72) Inventor: **Hetherington, Phillip Alan**  
**Vancouver, British Columbia**  
**V6B 2K4 (CA)**

(74) Representative: **Roberts, Gwilym Vaughan et al**  
**Kilburn & Strode LLP**  
**20 Red Lion Street**  
**London WC1R 4PJ (GB)**

(54) **Maintaining spatial stability utilizing common gain coefficient**

(57) In a system and method for maintaining the spatial stability of a sound field a background noise estimate may be estimated for each of a first signal and a second signal. A first gain coefficient may be calculated responsive to the first audio signal and the background noise estimate of the first audio signal. A second gain coefficient may be calculated responsive to the second signal and the background noise estimate of the second signal.

The gain coefficients may be calculated using one or more gain coefficient calculators. A common gain coefficient may be selected from one of the first gain coefficient and the second gain coefficient. The selected common gain coefficient may be one that results in a least amount of audio signal modification and may be applied to each of the first signal and the second signal.



**Description****BACKGROUND****1. Technical Field**

**[0001]** The present disclosure relates to the field of processing sound fields. In particular, to a system and method for maintaining the spatial stability of a sound field utilizing a common gain coefficient.

**2. Related Art**

**[0002]** Stereo and multichannel audio configurations may be used for processing a sound field that is a spatial representation of an audible environment. The processed sound field may be used to reproduce the audible environment using audio transducers.

**[0003]** Many computing devices may have multiple integrated microphones used for recording an audible environment associated with the computing device and communicating with other users. Computing devices typically use multiple microphones to improve noise performance with noise suppression processes. The noise suppression processes may result in the reduction or loss of spatial information. In many cases the noise suppression processing may result in a single, or mono, output signal that has no spatial information.

**BRIEF DESCRIPTION OF DRAWINGS**

**[0004]** The system may be better understood with reference to the following drawings and description. The components in the figures are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the disclosure. Moreover, in the figures, like referenced numerals designate corresponding parts throughout the different views.

**[0005]** Other systems, methods, features and advantages will be, or will become, apparent to one with skill in the art upon examination of the following figures and detailed description. It is intended that all such additional systems, methods, features and advantages be included with this description and be protected by the following claims.

**[0006]** Fig. 1 is a schematic representation of a system for maintaining the spatial stability of a sound field.

**[0007]** Fig. 2 is a further schematic representation of a system for maintaining the spatial stability of the sound field.

**[0008]** Fig. 3 is flow diagram representing a method for maintaining the spatial stability of the sound field.

**DETAILED DESCRIPTION**

**[0009]** In a system and method for maintaining the spatial stability of a sound field a background noise estimate may be estimated for each of a first audio signal and a second audio signal. A first gain coefficient may be calculated responsive to the first audio signal and the background noise estimate of the first audio signal. Gain coefficients may be calculated using a gain coefficient calculator that may include one or more of an automatic gain controller, a noise suppressor and an echo canceller. A second gain coefficient may be calculated responsive to the second audio signal and the background noise estimate of the second audio signal. The second gain coefficient may be calculated with one or more gain coefficient calculations similar to those applied to the first signal. A common gain coefficient may be selected from one of the first gain coefficient and the second gain coefficient. Selecting a common gain coefficient may comprise selecting a gain coefficient from the first gain coefficient and the second gain coefficient that will result in a least amount of audio signal modification. The selected common gain coefficient may be applied to each of the first audio signal and the second audio signal.

**[0010]** Figure 1 is a schematic representation of a system for maintaining the spatial stability of a sound field. Two or more inputs 102, or audio inputs 102, may receive the sound field. Stereo and multichannel input configurations may be utilized for processing the sound field that is a spatial representation of an audible environment associated. The audible environment may be associated with microphones on a local computing device or a remote computing device. The remote computing device may transmit audio signals 120 to the local computing device that may utilize the received audio signals 120 as inputs 102. Many audible environments associated with the inputs 102 may include undesirable content that may be mitigated by processing the received sound field. Microphones that are arranged in a far field configuration typically receive more undesirable content (a.k.a. noise) than microphones in a near field configuration. Far field configurations may include, for example, a hands free phone, a conference phone and microphones embedded into an automobile. Far field configurations are capable of receiving a sound field that represents the spatial environment associated with the microphones. Near field configurations typically place the microphone in close proximity to a user. Undesirable content may be mitigated in both near and far field configurations by processing the received sound field.

**[0011]** Processing that may mitigate undesirable content received in the sound field may include a gain coefficient

calculator 106. The gain coefficient calculator 106 may comprise one or more of a noise suppressor 110 and an echo canceller 112. The echo canceller 112, noise suppressor 110 and other audio processing processes may calculate one or more gain coefficients. Each respective gain coefficient may be applied individually or a composite gain coefficient may be applied to process the sound field using a gain coefficient applier 118.

**[0012]** The echo canceller 112 mitigates echoes caused by signal feedback between two or more communication devices. Signal feedback occurs when an audio transducer on a first communication device reproduces the signal received from a second communication device and subsequently the microphones on the first communication device recapture the reproduced signal. The recaptured signal may be transmitted to the second communication device where the recaptured signal may be perceived as an echo of the previously transmitted signal. The echo canceller 112 may detect when the signal has been recaptured and attempt to suppress the recaptured signal. Many different types of echo cancellers 112 may mitigate echoes by calculating one or more gain coefficients that, when applied to the signals received by the microphone, suppress the echoes. In one example implementation, the echo suppressor 112 may calculate gain coefficients using a coherence calculation between near and far signals disclosed in U.S. Patent No. 8,036,879, which is incorporated herein by reference, except that in the event of any inconsistent disclosure or definition from the present specification, the disclosure or definition herein shall be deemed to prevail.

**[0013]** When the microphone, or source of input 102, and an audio transducer are close in proximity, the echo canceller 112 may determine that a large amount of suppression may mitigate the signal produced by the audio transducer from dominating, or coupling with, the microphone. The echo canceller 112 may calculate large gain coefficients to mitigate the coupling. The large gain coefficients may result in a gating effect where the communication device effectively supports only half duplex communication. Half duplex communication may occur when the communication channel allows for reliable communication from alternatively either the far side or near side but not both simultaneously. The large gain coefficients may suppress the coupling but may also suppress all content, including desired voice content resulting in half duplex communication.

**[0014]** Background noise is another type of undesirable signal content that may be mitigated by processing the input 102. Many different types of noise suppressor 110 techniques may mitigate background noise. An exemplary noise suppressor 110 is a recursive Wiener filter. The Wiener suppression gain  $G_{i,k}$ , or gain coefficient, is defined as:

$$G_{i,k} = \frac{S\hat{N}R_{priori,i,k}}{S\hat{N}R_{priori,i,k} + 1}. \quad (1)$$

**[0015]** Where  $S\hat{N}R_{priori,i,k}$  is the a priori SNR estimate and is calculated recursively by:

$$S\hat{N}R_{priori,i,k} = G_{i-1,k} S\hat{N}R_{post,i,k} - 1. \quad (2)$$

**[0016]** Where  $S\hat{N}R_{post,i,k}$  is the a posteriori SNR estimate given by:

$$S\hat{N}R_{post,i,k} = \frac{|Y_{i,k}|^2}{|\hat{N}_{i,k}|^2}. \quad (3)$$

**[0017]** Where  $|\hat{N}_{i,k}|$  is a background noise estimate. A background noise estimator 104 may estimate the background noise estimate. In one example implementation, the background noise estimate, or signal values, may be calculated using the background noise estimation techniques disclosed in U.S. Patent No. 7,844,453, which is incorporated herein by reference, except that in the event of any inconsistent disclosure or definition from the present specification, the disclosure or definition herein shall be deemed to prevail. In other implementations, alternative background noise estimation techniques may be used, such as, for example, a noise power estimation technique based on minimum statistics.

**[0018]** An automatic gain controller 108 may calculate gain coefficients that may mitigate changing energy levels of the desired signal content. For example, the energy level of a user speaking into the microphone may change over time as the microphone may change position relative to the user. The gain coefficients calculated by the automatic gain controller 108 may mitigate the perception of the microphone changing position when applied to the input 102. In one example implementation, the automatic gain controller 108 may calculate gain coefficients using the gain controller

techniques disclosed in U.S. Patent No. 8,116,485, which is incorporated herein by reference, except that in the event of any inconsistent disclosure or definition from the present specification, the disclosure or definition herein shall be deemed to prevail.

**[0019]** The gain coefficient calculators 106 including the automatic gain controller 108, the noise suppressor 110 and the echo canceller 112 described above may be responsive to a background noise estimate generated by the background noise estimator 104. The automatic gain controller 108 may utilize the background noise estimate to calculate gain coefficients that may be adjusted when signal energy exceeds the background noise estimate resulting in less background noise being amplified. The echo canceller 112 may utilize the background noise estimate to calculate gain coefficients when the echoes exceed the background noise estimate by a threshold. The background noise estimator 104 may calculate a background noise estimate for each input 102.

**[0020]** When the inputs 102 are generated by physically separated microphones, or when two or more inputs 102 do not contain identical signal content, the background noise estimator 104 and the gain coefficient calculator 106 may calculate different background noise estimates and gain coefficients. Differences between the first gain coefficients applied to the first input 102 and second gain coefficients applied to the second input 102 may cause a distortion in the spatial image when reproduced in the output 116. Different gain coefficients applied to the first input 102 and the second input 102 may result in a shifting spatial image that may be distracting to a listener.

**[0021]** A common gain coefficient selector 114 may mitigate some distortion in the spatial image by selecting a first gain coefficient or a second gain coefficient that may be applied to both the first input 102 and the second input 102. Applying the same gain coefficient to all inputs 102 may mitigate distortions in the spatial image. The largest or the smallest gain coefficient may be selected. Alternatively, a combination of the gain coefficients may be calculated. For example, when the first input 102 contains a larger echo than the second input 102, an average gain coefficient may be calculated to perceptibly remove the larger echo. In another alternative, the common gain coefficient selector 114 may select the gain coefficient that will result in a least amount of audio signal modification. For example, the echo canceller 112 may calculate a first gain coefficient for application to the first input 102 that is larger than a second gain coefficient for application to the second input 102. The larger gain coefficient may result in a lower energy signal to reduce the amount of echo in the first input 102. The common gain coefficient selector 114 may select the smaller gain coefficient for application to both the first input 102 and the second input 102 resulting in less echo suppression while mitigating distortions in the spatial image. The common gain coefficient selector 114 may utilize two or more inputs 102 where the selected gain coefficient is applied to the two or more input 102. For example, the common gain coefficient selector 114 in addition to the first input 102 and the second input 102 may process a third input 102 and their respective calculated gain coefficients.

**[0022]** The common gain coefficient selector 114 may not apply to all types of gain coefficient calculators 106. For example, some types of noise reduction processes may require additional processing to mitigate distortions in the spatial image. A first input 102 that contains significant wind noise may have wind noise reduction applied while the second input 102 has no wind noise reduction applied. The common gain coefficient selector 114 may not mitigate the wind noise that may be more distorting than a shift in the spatial image. In one implementation, wind noise suppression gains (a.k.a. gain coefficients) may be calculated using the system for suppressing wind noise disclosed in U.S. Patent No. 7,885,420, which is incorporated herein by reference, except that in the event of any inconsistent disclosure or definition from the present specification, the disclosure or definition herein shall be deemed to prevail. In another example, when the microphone and audio transducer are coupled as described above resulting in a gating effect, the common gain coefficient selector 114 may not be utilized because the audible artifacts associated with the coupling are perceptibly more distracting than distorting the spatial image.

**[0023]** A subband filter may process the received input 102 to extract frequency information. The subband filter may be accomplished by various methods, such as a Fast Fourier Transform (FFT), critical filter bank, octave filter band, or one-third octave filter bank. Alternatively, the subband analysis may include a time-based filter bank. The time-based filter bank may be composed of a bank of overlapping bandpass filters, where the center frequencies have non-linear spacing such as octave, 3<sup>rd</sup> octave, bark, mel, or other spacing techniques. The one or more gain coefficients may be calculated for each frequency bin or band of the subband filter. The gain coefficients and the selected gain coefficients may be filtered, or smoothed, over time and/or frequency.

**[0024]** Figure 3 is flow diagram representing a method for maintaining the spatial stability of the sound field. The method 300 may be, for example, implemented using the systems 100 and 200 described herein with reference to Figures 1 and 2. The method 300 may include the following acts. Estimating a background noise estimate for each of a first audio signal and a second audio signal 302. Calculating a first gain coefficient responsive to the first audio signal and the background noise estimate of the first audio signal 304. The gain coefficients may be calculated using a gain coefficient calculator that may include one or more of an automatic gain controller, a noise suppressor and an echo canceller. Calculating a second gain coefficient responsive to the second audio signal and the background noise estimate of the second audio signal 306. The second gain coefficient may be calculated with one or more of the same gain coefficient calculations similar to those applied to the first signal. Selecting a common gain coefficient from one of the

first gain coefficient and the second gain coefficient 308. Selecting the common gain coefficient may comprise selecting a gain coefficient from the first gain coefficient and the second gain coefficient that will result in a least amount of audio signal modification. Applying the selected common gain coefficient to each of the first audio signal and the second audio signal 310.

**[0025]** Figure 2 is a further schematic representation of a system for maintaining the spatial stability of a sound field. The system 200 comprises a processor 202, memory 204 (the contents of which are accessible by the processor 202) and an I/O interface 206. The memory 204 may store instructions which when executed using the processor 202 may cause the system 200 to render the functionality associated with the system for maintaining the spatial stability of the sound field as described herein. In addition the memory 204 may store instructions which when executed using the processor 202 may cause the system 200 to render the functionality associated with the background noise estimator 104, the gain coefficient calculator 106, the automatic gain controller 108, the noise suppressor 110, the echo canceller 112, the common gain coefficient selector 114 and the gain coefficient applier 118 described herein. In addition, data structure, temporary variables and other information may be stored in data storage 208.

**[0026]** The processor 202 may comprise a single processor or multiple processors that may be disposed on a single chip, on multiple devices or distributed over more than one system. The processor 202 may be hardware that executes computer executable instructions or computer code embodied in the memory 204 or in other memory to perform one or more features of the system. The processor 202 may include a general purpose processor, a central processing unit (CPU), a graphics processing unit (GPU), an application specific integrated circuit (ASIC), a digital signal processor (DSP), a field programmable gate array (FPGA), a digital circuit, an analog circuit, a microcontroller, any other type of processor, or any combination thereof.

**[0027]** The memory 204 may comprise a device for storing and retrieving data, processor executable instructions, or any combination thereof. The memory 204 may include non-volatile and/or volatile memory, such as a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM), or a flash memory. The memory 204 may comprise a single device or multiple devices that may be disposed on one or more dedicated memory devices or on a processor or other similar device. Alternatively or in addition, the memory 204 may include an optical, magnetic (hard-drive) or any other form of data storage device.

**[0028]** The memory 204 may store computer code, such as the background noise estimator 104, the gain coefficient calculator 106, the automatic gain controller 108, the noise suppressor 110, the echo canceller 112, the common gain coefficient selector 114 and the gain coefficient applier 118 as described herein. The computer code may include instructions executable with the processor 202. The computer code may be written in any computer language, such as C, C++, assembly language, channel program code, and/or any combination of computer languages. The memory 204 may store information in data structures including, for example, the gain coefficients.

**[0029]** The I/O interface 206 may be used to connect devices such as, for example, the input 102 and output 116 to other components of the system 200.

**[0030]** All of the disclosure, regardless of the particular implementation described, is exemplary in nature, rather than limiting. The systems 100 and 200 may include more, fewer, or different components than illustrated in Figures 1 and 2. Furthermore, each one of the components of systems 100 and 200 may include more, fewer, or different elements than is illustrated in Figures 1 and 2. Flags, data, databases, tables, entities, and other data structures may be separately stored and managed, may be incorporated into a single memory or database, may be distributed, or may be logically and physically organized in many different ways. The components may operate independently or be part of a same program or hardware. The components may be resident on separate hardware, such as separate removable circuit boards, or share common hardware, such as a same memory and processor for implementing instructions from the memory. Programs may be parts of a single program, separate programs, or distributed across several memories and processors.

**[0031]** The functions, acts or tasks illustrated in the figures or described may be executed in response to one or more sets of logic or instructions stored in or on computer readable media. The functions, acts or tasks are independent of the particular type of instructions set, storage media, processor or processing strategy and may be performed by software, hardware, integrated circuits, firmware, micro code and the like, operating alone or in combination. Likewise, processing strategies may include multiprocessing, multitasking, parallel processing, distributed processing, and/or any other type of processing. In one embodiment, the instructions are stored on a removable media device for reading by local or remote systems. In other embodiments, the logic or instructions are stored in a remote location for transfer through a computer network or over telephone lines. In yet other embodiments, the logic or instructions may be stored within a given computer such as, for example, a CPU.

**[0032]** While various embodiments of the system and method for maintaining the spatial stability of a sound field have been described, it will be apparent to those of ordinary skill in the art that many more embodiments and implementations are possible within the scope of the present invention. Accordingly, the invention is not to be restricted except in light of the attached claims and their equivalents.

## Claims

1. A computer implemented method for maintaining the spatial stability of a sound field comprising:

5       estimating (302) a background noise estimate for each of a first audio signal (120) and a second audio signal (120);  
       calculating (304) a first gain coefficient responsive to the first audio signal (120) and the background noise estimate of the first audio signal (120);  
       calculating (306) a second gain coefficient responsive to the second audio (120) signal and the background  
 10       noise estimate of the second audio signal (120);  
       selecting (308) a common gain coefficient from the first gain coefficient and the second gain coefficient; and  
       applying (310) the selected common gain coefficient to each of the first audio (120) signal and the second audio signal (120).

- 15       2. The method for maintaining the spatial stability of a sound field of claim 1, where selecting the common gain coefficient comprises selecting a gain coefficient, from the first gain coefficient and the second gain coefficient, that will result in a least amount of audio signal (120) modification.

- 20       3. The method for maintaining the spatial stability of a sound field of claims 1 to 2, where selecting the common gain coefficient comprises selecting a gain coefficient, from the first gain coefficient and the second gain coefficient, that will result in a greatest amount of audio signal (120) modification.

- 25       4. The method for maintaining the spatial stability of a sound field of claims 1 to 3, where selecting the common gain coefficient comprises averaging a gain coefficient, from the first gain coefficient and the second gain coefficient, that will result in a combined average amount of audio signal (120) modification.

5. The method for maintaining the spatial stability of a sound field of claims 1 to 4, further comprising:

30       estimating a background noise estimate for a third audio signal (120);  
       calculating a third gain coefficient responsive to the third audio signal (120) and the background noise estimate of the third audio signal (120);  
       selecting the common gain coefficient from the first gain coefficient, the second gain coefficient and the third gain coefficient; and  
       applying the selected common gain coefficient to each of the first audio signal (120), the second audio signal  
 35       (120) and the third audio signal (120).

- 40       6. The method for maintaining the spatial stability of a sound field of claims 1 to 5, where calculating each of the first gain coefficient and the second gain coefficient includes one or more of a noise suppression calculation, an echo cancellation calculation and an automatic gain control calculation.

7. The method for maintaining the spatial stability of a sound field of claims 1 to 6, further comprising generating a set of sub-bands for each of the first audio signal (120) and the second audio (120) signal using a subband filter or a Fast Fourier Transform.

- 45       8. The method for maintaining the spatial stability of a sound field of claims 1 to 7, further comprising generating a set of sub-bands for each of the first audio signal (120) and the second audio signal (120) according to a critical, octave, mel, or bark band spacing technique.

- 50       9. A system for maintaining spatial stability of a sound field, the system comprising:

      a processor (202); and  
       a memory (204) coupled to the processor (202) containing instructions,  
       executable by the processor (202), for performing the instructions executing the steps of any of method claims  
 55       1 to 8.

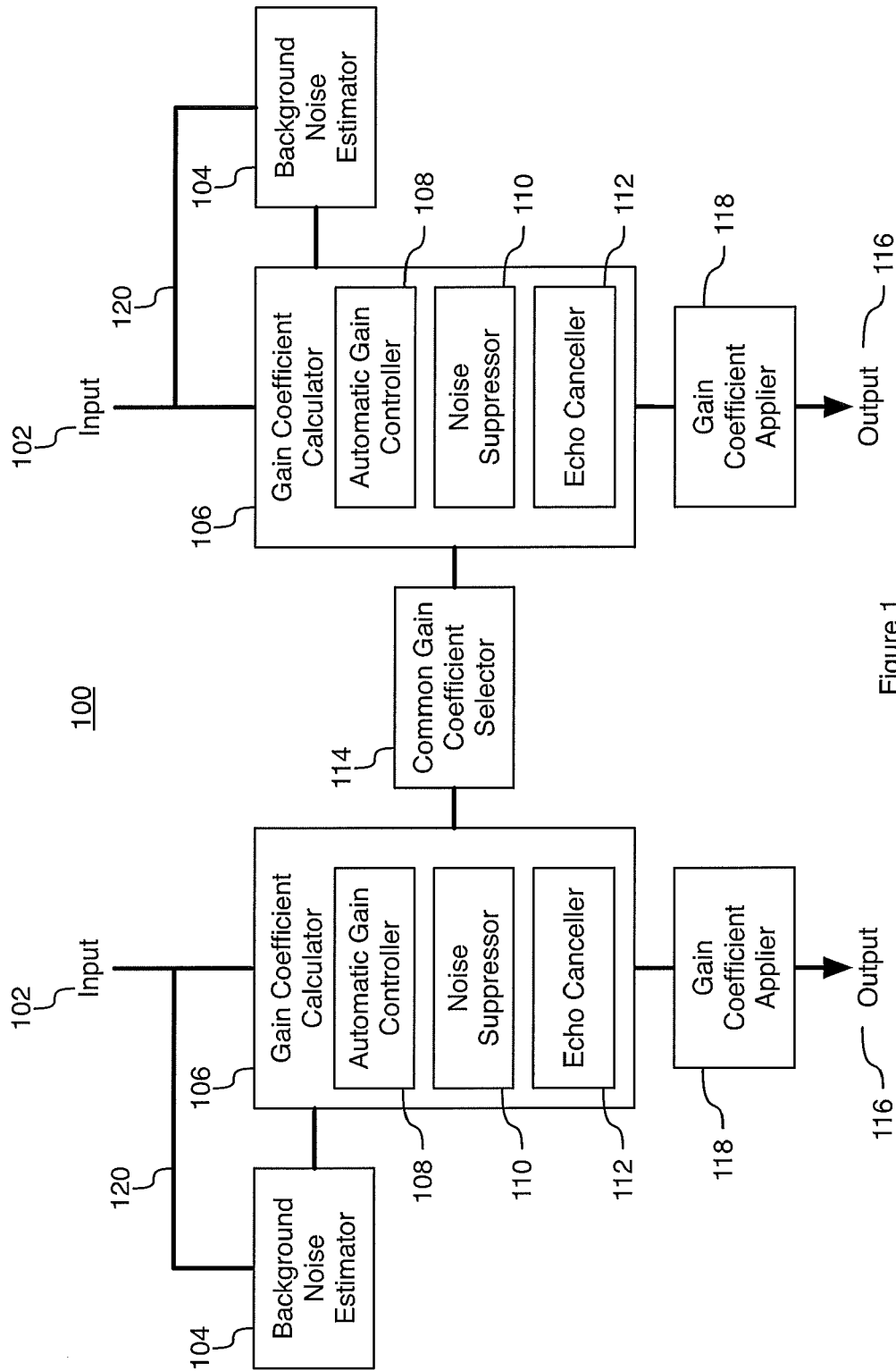


Figure 1

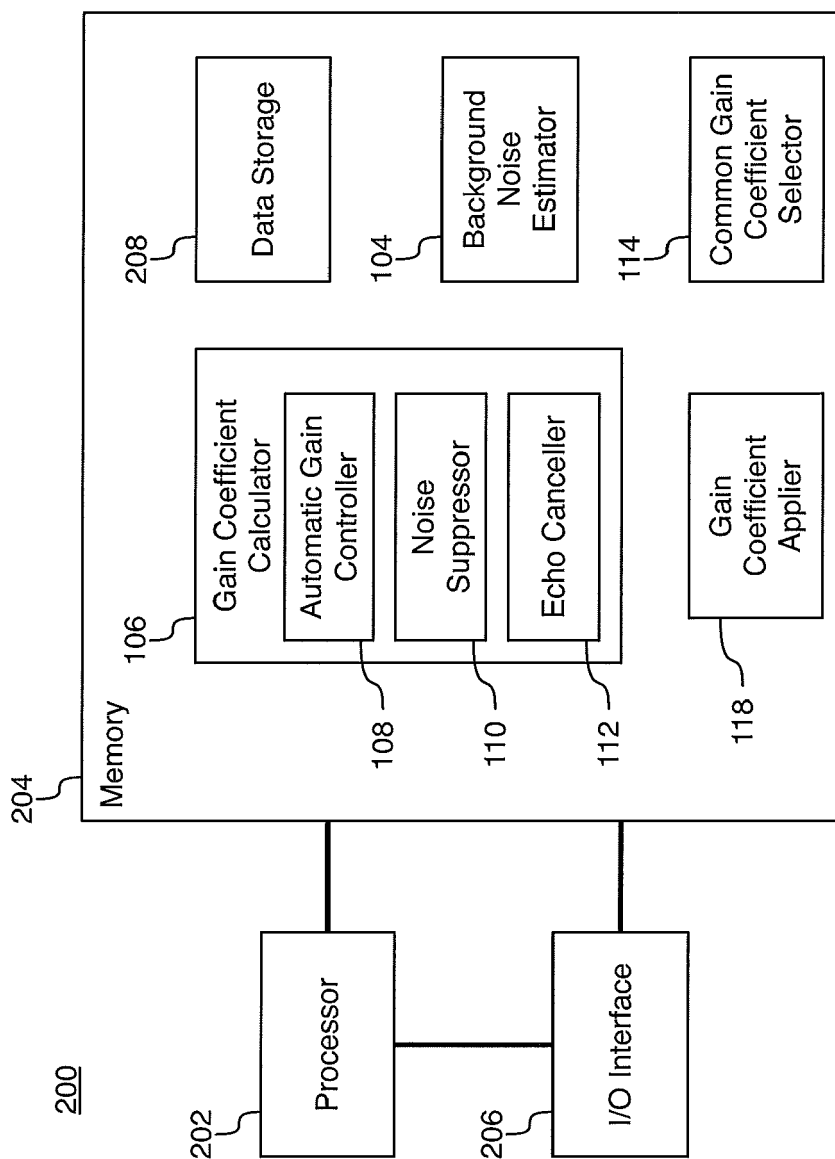


Figure 2



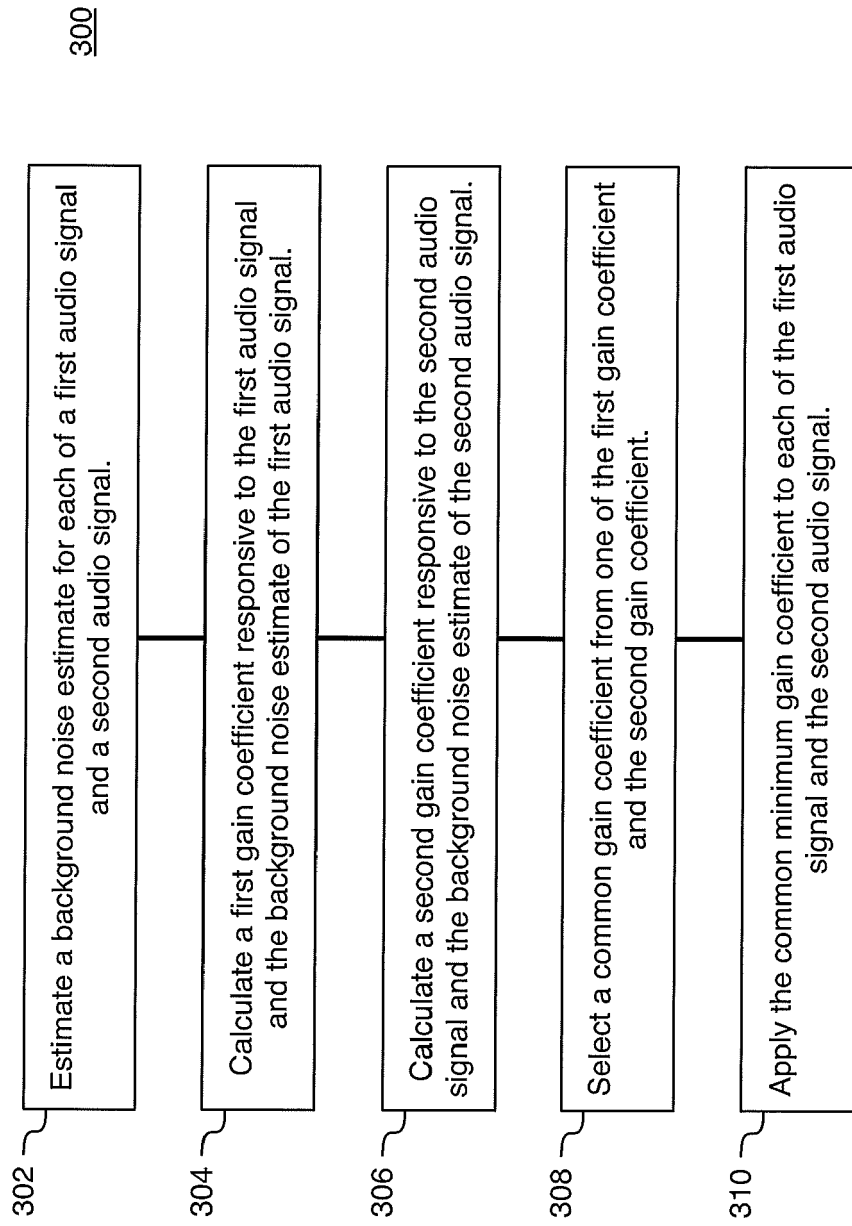


Figure 3



## EUROPEAN SEARCH REPORT

 Application Number  
 EP 13 15 3062

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	WO 2009/096958 A1 (AGERE SYSTEMS INC [US]; GAENSLER TOMAS FRITZ [US]; JOHANSON JAMES A [U] 6 August 2009 (2009-08-06) * paragraphs [0024], [0027], [0033] - [0034], [0037] - [0044]; figures 1,4,5 * -----	1-9	INV. G10L21/0208
X	RÉGINE LE BOUQUIN JEANN GBP ESLE BOUQUIN JEANNES ET AL: "Combined Noise and Echo Reduction in Hands-Free Systems: A Survey", IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, vol. 9, no. 8, 30 November 2001 (2001-11-30), XP011054146, ISSN: 1063-6676 * page 817, right-hand column, line 21 - page 818, left-hand column, line 15; figure 12 * -----	1,6-9	TECHNICAL FIELDS SEARCHED (IPC)  G10L
The present search report has been drawn up for all claims			
Place of search <b>Munich</b>		Date of completion of the search <b>12 April 2013</b>	Examiner <b>Dobler, Ervin</b>
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 13 15 3062

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.

12-04-2013

10

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2009096958 A1	06-08-2009	NONE	
-----			

15

20

25

30

35

40

45

50

EPO FORM P0459

55

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

**REFERENCES CITED IN THE DESCRIPTION**

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

**Patent documents cited in the description**

- US 8036879 B [0012]
- US 7844453 B [0017]
- US 8116485 B [0018]
- US 7885420 B [0022]