

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
12 February 2004 (12.02.2004)

PCT

(10) International Publication Number
WO 2004/014105 A1

(51) International Patent Classification⁷: H04S 1/00, G10H 1/36

(21) International Application Number: PCT/IB2003/003389

(22) International Filing Date: 31 July 2003 (31.07.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data: 02078147.2 31 July 2002 (31.07.2002) EP

(71) Applicant: KONINKLIJKE PHILIPS ELECTRONICS N.V. [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors: RODENAS, Josep, A.; c/o Prof . Holstlaan 6, NL-5656 AA Eindhoven (NL). IRWAN, Roy; c/o Prof

. Holstlaan 6, NL-5656 AA Eindhoven (NL). AARTS, Ronaldus, M.; c/o Prof . Holstlaan 6, NL-5656 AA Eindhoven (NL).

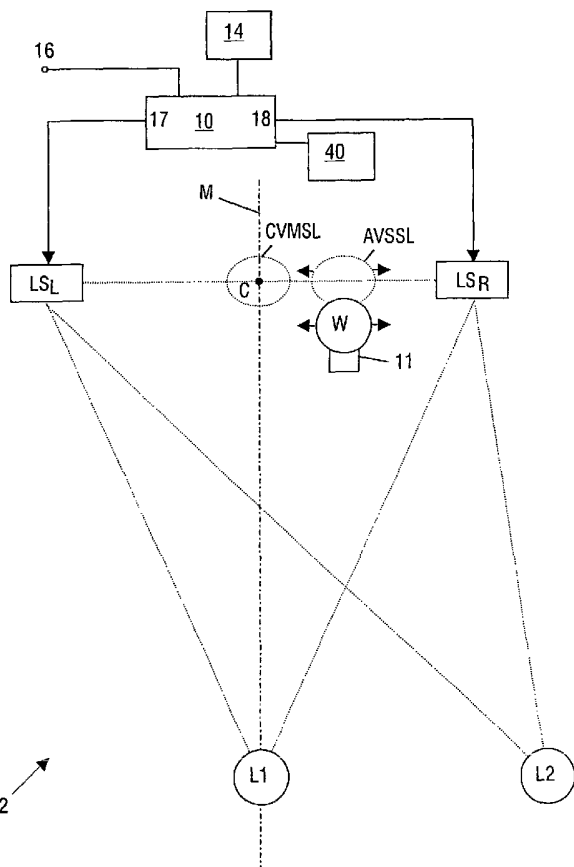
(74) Agent: GROENENDAAL, Antonius, W., M.; Philips Intellectual Property & Standards, Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,

[Continued on next page]

(54) Title: AUDIO PROCESSING SYSTEM



(57) Abstract: A sound producing system (1) comprises two loudspeakers (LS_L; LS_R) arranged next to each other on opposite sides of a median plane (M). Each loudspeaker has a radiation characteristics (R_{2L}; R_{2R}) designed such that a relatively large sweet spot is achieved. The sound producing system further comprises an audio processing system (10) having a microphone input (16) and loudspeaker outputs (17; 18), and adapted to generate loudspeaker drive signal (SD_L; SD_R) having a certain balance ratio (ρ_{LR}). Source location detection means (40) are associated with the audio processing system (10) for generating a location signal representing the actual source location of a microphone (11). The audio processing system (10) is responsive to the location signal received from the source location detection means (40) to amend the setting of said balance ratio (ρ_{LR}).

WO 2004/014105 A1



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

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

Published:

— *with international search report*

Audio processing system

The present invention relates in general to an audio processing system, especially suitable for presenting music and/or songs to an audience. More specifically, the present invention relates to an audio processing system suitable for karaoke.

In the following, the phrase "driver" will be used for a device capable of converting electric input power to output sound waves. In some conventions, such device is also indicated as speaker or loudspeaker, but in the context of the present invention the phrase "loudspeaker" will be used for an assembly comprising a housing or cabinet and one or more drives mounted in said housing, while the phrase "speaker" will be used for a person, such as for instance a person speaking, or singing, or playing an instrument; however, in order to try to avoid misconception, usually the phrase "singer" will be used for such person.

Stereo audio processing systems for presenting sound or song to an audience are generally known. Basically, such system comprises two or more loudspeakers driven by a stereo amplifier device, which may receive a conventional audio signal from a conventional source, such as a recording on CD. As is commonly known, stereo systems have a problem relating to the fact that listeners are capable of perceiving a direction from which sounds originates. In the case of two loudspeakers, a listener will perceive the sound as originating from a virtual source having a virtual source location which depends on the location of the listener. In a symmetric set-up, a listener positioned in a median plane with respect to said loudspeakers will perceive the virtual source location at a central location between said two loudspeakers. However, a listener positioned outside said median plane will perceive the virtual source location as substantially coinciding with the location of the closest loudspeaker.

The problem becomes somewhat more complicated if the audio source is visible to the listeners. This applies in the case of an "imaged" audio source such as a television screen, and also in the case of a "true" audio source such as a person carrying a microphone, who may for instance be a speaker or a singer or a musician, and who is free to move on a stage. Listeners will perceive the microphone-generated sound as originating from a fixed virtual source location, which will be a fixed location in space, for instance coinciding with the center between said two loudspeakers or substantially coinciding with the location of

the closest loudspeaker. However, the speaker or singer may be moving on a stage or the like, so that in general the physical location of the speaker or singer does not correspond to the virtual source location as perceived by a listener. For such listener, the experience of seeing a person in one location and perceiving his voice originating from another location is a strange experience, lacking reality.

It is an objective of the present invention to solve or at least mitigate this problem.

The problem becomes even more complicated if a plurality of independently moving audio sources is visible to the listeners, for instance a group of singers and/or musicians. Now, all sound seems to originate from one fixed virtual source location, whereas the listener can see different persons at different locations.

It is an objective of the present invention to also solve or at least mitigate this problem.

According to an important aspect of the present invention, an audio processing system comprises actual source location detection means for generating a signal indicative of the actual location of an audio source such as a speaker, and processing means responsive to this detecting means to adapt the balance of the driving means for the loudspeakers such that the virtual source location is shifted such as to substantially correspond to the speaker location as detected.

In one specific embodiment, the actual source location detecting means comprise a signal transmitter associated with the microphone or carried by the speaker/singer. In this embodiment, the detection means may further comprise a system of receivers adapted to receive the signal emitted by the sender, and a processor adapted to combine the receiver signals such as to determine the sender location (for instance by triangulation method).

In another, more sophisticated embodiment, the audio processing system is adapted to generate a microphone-bearing signal, preferably having a frequency outside the audible range, which microphone-bearing signal is picked-up by the microphone and sent back to the audio processing system, which is capable of calculating the distance from the microphone to each loudspeaker and hence to calculate the location of the microphone in space with respect to the coordinate system as defined by the loudspeakers.

The problem as explained above already exists in the case of a speaker or singer or musician speaking or singing or playing without background music. The problem is even more complicated in the case of a singer or musician who is accompanied by

(background) music. In such case, a listener should ideally perceive two virtual source locations, one associated with the singer and one associated with the music. This music may originate from a live instrument or live orchestra, having a fixed location on stage, but the music may also originate from a sound carrier such as an audio disc, such as for instance in the case of karaoke, in which case a virtual music source location is associated with the recording. In conventional audio processing systems, no measures are taken to assure a correct positioning of said two virtual source locations. Therefore, according to a further important aspect, the present invention provides an audio processing system capable of producing audio with one set of loudspeakers in such a way that two or even more virtual source locations are generated, i.e. (at least) one virtual music source location and (at least) one virtual singer source location, wherein the virtual music source location is a fixed location independent from the actual singer location, and wherein the virtual singer source location moves with the position of the singer as he moves around the stage. In an embodiment, an audio system comprises at least two sound processing channels, a first frequency channel substantially comprising music signals and a second channel substantially comprising microphone signals. Audio signals within the first channel are treated such as to result in a fixed virtual source location. In contrast, audio signals within the second channel are treated such as to result in a variable source location corresponding to the variable singer location.

Another problem relates to the speaker or singer himself. Since he can hear his own voice being reproduced by the loudspeakers, he can be considered as being one of the listeners, with the important difference that he will generally be located closer to the loudspeakers than the actual audience. He also will perceive a virtual source location, but due to his being close to the loudspeakers, the virtual source location as perceived by the singer will typically coincide with the location of the closest loudspeaker.

The present invention also aims to provide a solution to this problem.

In one approach, the balance of the loudspeakers is adapted such that the virtual singer source location as perceived by the singer is a fixed location, for instance corresponding to the center C between the two loudspeakers, or, alternatively such that the virtual singer source location as perceived by the singer is a variable location moving along with the moving singer.

In another approach, the singer wears a set of earphones, and the balance of signal supplied to the set of earphones is set such that the virtual singer source location as perceived by the singer is a fixed location halfway the two earphones.

These and other aspects, features and advantages of the present invention will be further explained by the following description with reference to the drawings, in which same reference numerals indicate same or similar parts, and in which:

Fig. 1 is a schematic top view of an arrangement of loudspeakers and listeners;

Fig. 2 is a schematic top view of an arrangement of loudspeakers and listeners, for illustrating an adaptable virtual source location according to the invention;

Figs. 3A and 3B are schematical block diagrams illustrating microphone location detection means;

Fig. 4 is a schematical block diagram illustrating an audio processing system; and

Fig. 5 is a schematical block diagram illustrating another embodiment of an audio processing system.

The present invention will first be explained for the case of one speaker/singer/musician; later, the case of a plurality of speakers/singers/musicians will be discussed.

Reference is made to Fig. 1, which schematically shows a sound producing system 1 comprising an arrangement of two loudspeakers LS_L and LS_R , arranged next to each other, at a certain mutual distance. Herein, each loudspeaker may comprise one or more individual drivers. A central location halfway between said two loudspeakers LS_L and LS_R is indicated at C. A median plane between said two loudspeakers LS_L and LS_R is indicated at M. More specifically, the two loudspeakers LS_L and LS_R are placed in a symmetrical arrangement with respect to this median plane M. The subscripts L and R denote left and right, as seen by a listener. The figure shows a first listener L1 located in a position coinciding with said median plane M, and a second listener L2 located besides the first listener L1, at a certain distance from said median plane M.

The figure also shows radiation diagrams $R1_L$ and $R1_R$ of said two loudspeakers LS_L and LS_R , respectively. A radiation diagram indicates the relative intensity of sound generated by the corresponding loudspeaker into a certain direction, assuming that all sound originates from one point S_L , S_R , respectively, which point is taken as origin of a polar coordinate system. For instance, with respect to the left-hand loudspeaker LS_L , a line

through said point S_L , parallel to said median plane M , is taken as X-axis in this coordinate system. A position A in front of the left loudspeaker LS_L is defined in polar coordinates r and ϕ , wherein r is the distance $|A-S_L|$ and wherein ϕ is the angle between the X-axis and the line AS_L .

The intensity of sound, generated by a loudspeaker into a direction ϕ , is indicated by the length of a line piece from said one originating point to the intersection with the corresponding radiation diagram. For instance, the relative intensity of sound into the direction of the first listener L1 is indicated by the length $|S_L-B_L|$. For calculating the absolute intensity of sound received at the location of the first listener L1, it may be assumed that the sound intensity decreases with the distance in a predetermined manner, typically with the inverse square of the distance.

The shape of such radiation diagram represents a characteristic of the loudspeaker in question. Normally, such diagram is symmetric with respect to $\phi=0$. By way of example, for a case where the loudspeakers radiate in a uniform manner, the radiation diagrams $R1_L$ and $R1_R$ constitute a (part of a) circle.

As is well known, when the two loudspeakers produce the same (mono) sound, a listener will perceive the sound as originating from a virtual source having a virtual source location VSL which depends on the location of the listener. Assuming that the two loudspeaker systems LS_L and LS_R have uniform radiation diagrams $R1_L$ and $R1_R$, and that they are driven in a symmetrical way, the first listener L1 will perceive the virtual source location VSL at said central location C, whereas the second listener L2 will perceive the virtual source location as substantially coinciding with the location of the closest loudspeaker, i.e. the right-hand loudspeaker LS_R in Fig. 1.

In the case of stereo music played by such audio system, the first listener L1 will perceive good stereo quality, involving a good separation between left-hand channel sound and right-hand channel sound, whereas the second listener L2 will perceive all sound as originating from the right-hand loudspeaker LS_R , which seriously affects or even eliminates the stereo perception.

This is mainly caused by two effects. First, the listener L2 receives sound from the closest loudspeaker LS_R earlier than the sound from the more remote loudspeaker LS_L . Second, the listener L2 receives sound from the closest loudspeaker LS_R with larger intensity than the sound from the more remote loudspeaker LS_L .

The area in space where good stereo quality is perceived is indicated as the "sweet spot". In the conventional setup described above, the sweet spot will more or less

coincide with the median plane M. In an improved and preferred setup, the sweet spot is broadened by using loudspeaker having a non-uniform radiation characteristic, such that the sound intensity radiated towards the median plane is larger than the sound intensity radiated in the direction $\phi=0$. This is schematically illustrated in Fig. 1 by radiation diagrams R_{2L} and R_{2R} , respectively.

Again, the two loudspeakers are arranged symmetrically, while further said two non-uniform radiation characteristics are mutually mirror-symmetrical, so that the situation does not change for listeners located along the median plane M.

Again, the second listener L2 receives sound from the closest loudspeaker LS_R earlier than the sound from the more remote loudspeaker LS_L . However, the second listener L2 receives sound from the closest loudspeaker LS_R with relatively less intensity as compared to the intensity of sound received from the more remote loudspeaker LS_L . This intensity difference can be more than 10 dB.

As is known per se, this intensity difference at the location of the second listener L2 can counter-act the effect of the earlier arriving sound wave from the closest loudspeaker, such that the listener L2 will also perceive a virtual source location at C. Thus, effectively, the extent of the area (sweet spot) where all listeners will perceive the same virtual source location has increased.

For a more detailed description of this known effect, indicated as time/intensity trading, reference is made to the book "Spatial hearing: The Psychophysics of Human Sound Localization" by J. Blauert, The MIT Press, 1983.

Fig. 1 also shows schematically a movable audio source, indicated at W, which source is free to move on a stage, as indicated by arrows in an Y-direction. By way of example, it will be assumed that this source W is a person, who may be speaking, singing, or playing his instrument. Sound generated by this person W is picked up by a microphone 11 carried by this person W. As mentioned above, the audio signal produced by the microphone is processed by the audio processing system, and sound is emitted from the loudspeakers LS_L and LS_R . In a conventional system, the microphone signal S_M is treated as "normal" mono signal, and is fed, after suitable amplification, to both loudspeakers LS_L and LS_R in equal intensities. As explained above, the listeners will perceive the microphone-generated sound as originating from a fixed virtual source location VSL, which will be a fixed location in space. However, the speaker or singer may be moving on a stage or the like, so that in general the physical location of the speaker or singer W does not correspond to the virtual source location VSL as perceived by a listener.

In the present invention, a balance ratio parameter ρ_{LR} will be defined as the ratio G_L/G_R , wherein G_L and G_R indicate the gain of the left-hand audio channel and the right-hand audio channel, respectively. More particularly, the balance ratio parameter ρ_{LR} may depend on signal frequency f . The balance ratio parameter ρ_{LR} may be specified for a certain frequency, which will be indicated as $\rho_{LR}(f) = G_L(f)/G_R(f)$. The balance ratio parameter ρ_{LR} may also be specified as a substantially constant value for all frequencies within a frequency range from a first frequency f_1 to a second frequency f_2 , which will be indicated as

$$\rho_{LR}[f_1;f_2] = G_L[f_1;f_2]/G_R[f_1;f_2].$$

In a normally balanced mode, the balance ratio $\rho_{LR} = 1$; this means in practice that, if a mono signal is received, this signal is applied to the left-hand loudspeaker and to the right-hand loudspeaker with equal amplification. As explained above, this will lead to a virtual source location VSL located at said center C, at least for listeners within the sweet spot. The present invention is partly based on the understanding that the virtual source location VSL can be shifted towards either loudspeaker by changing the balance ratio parameter ρ_{LR} .

In the following, with reference to Fig. 2, a sound production system 2 in accordance with the present invention will be described, which is capable of providing a movable virtual source location. First, the invention will be explained for the case of only one source. As will be explained, such sound production system 2 is particularly desirable in case the source of the audio is mobile and visible to the listeners. Such sound production system is already useful in the case of a relatively narrow sweet spot, but preferably it implements the widened sweet spot technology as described above.

As illustrated in Fig. 2, the sound production system 2 comprises an arrangement of two loudspeakers LS_L and LS_R , connected to loudspeaker outputs 17 and 18, respectively, of an audio processing system 10. Fig. 2 further indicates a person W, equipped with a microphone 11, who is free to move on a stage. The microphone 11 generates a microphone signal S_M which is transferred, either through a wire-coupling or a wireless coupling such as known per se, to a microphone input 16 of said audio processing system 10. Said audio processing system 10 receives and processes the microphone signal S_M and drives the loudspeakers LS_L and LS_R accordingly. As explained above, listeners will perceive the microphone sound as originating from a virtual source location. In accordance with the present invention, said audio processing system 10 provides for an adaptive virtual singer

source location AVSSL. More particularly, said audio processing system 10 is capable of generating sound in such a way that the virtual source location as perceived by a listener corresponds to the actual location of person W, as indicated in Fig. 2. Thus, if this person W is moving, indeed, said audio processing system 10 controls its output signals to the loudspeakers such that the adaptive virtual singer source location AVSSL is actually moving along with said person W.

To this end, the audio processing system 10 according to the present invention utilizes the effect that the virtual source location shifts towards the loudspeaker with the highest sound intensity. Thus, in the case depicted in Fig. 2, where the person W is located closer to the right-hand loudspeaker LS_R , the audio processing system 10 according to the present invention would generate its left-hand output drive signal at a lower magnitude than its right-hand output drive signal. More particularly, the audio processing system 10 according to the present invention is adapted to change the balance ratio ρ_{LR} in accordance with the actual location of person W.

It is noted that the audio processing system 10 may be used in combination with various types of loudspeakers. Preferably, however, the loudspeakers are of a type having an asymmetric radiation diagram R2L, R2R, as explained above with reference to Fig. 1, such that the shifting of the virtual source location AVSSL will be perceived by all listeners in relatively large sweet spot, while the shifted virtual source location will be approximately the same for all listeners in the increased sweet spot.

Ideally, the shifting of the virtual source location will be such that the shifted virtual source location substantially coincides with the actual source location of person W. To be able to do so, the audio processing system 10 needs to have information on the actual source location. To this end, the audio processing system 10 is provided with actual source location detection means 40.

In one embodiment, illustrated in Fig. 3A, such actual source location detection means 40 comprise a transmitter 41 carried by the person W or, preferably, associated with the microphone 11, and at least one receiver 42 coupled to the audio processing system 10 for sending a receiver signal S42 to the audio processing system 10.

The transmitter 41 is adapted to emit a predefined signal S41, which is received by the receiver 42. Said predefined signal S41 is such that the receiver 42, or alternatively the audio processing system 10, is capable of calculating the actual location of the transmitter 41.

There are various ways to implement this embodiment. For instance, the transmitter 41 may be associated with a GPS module (not shown), and the predefined signal S41 may actually communicate the GPS-coordinates to the audio processing system 10.

It is also possible that the system comprises an array of receivers 42 coupled to the audio processing system 10, and that the transmitter 41 is adapted to emit a pulsed signal, which may be a light signal or a radio signal but which preferably is a sound signal, more preferably an ultrasound signal. When the emitted signal S41 is received by the receivers 42, the times of arrival depend on the actual distances between the transmitter 41 and the respective receivers 42. Thus, on the basis of one emitted signal pulse S41, the audio processing system 10 will receive a plurality of receiver signals S42, the relative order and time differences being representative for the actual location of the transmitter 41 with respect to the receivers 42.

In another embodiment, illustrated in Fig. 3B, the audio processing system 10 is designed to generate predetermined microphone bearing signals S43 through the loudspeakers LS_L and LS_R , typically pulsed signals. Preferably, these location bearing signals S43 are generated within a frequency range inaudible to the human ear but within the capabilities of the loudspeakers and the microphone. The emitted signals S43 are picked up by the microphone 11, the respective times of arrival depending on the actual distances between the microphone 11 and the respective loudspeakers LS_L and LS_R . The signals S43 as picked up by the microphone 11 are converted into electrical signals S44 and, as part of the normal microphone signal S_M , are transmitted to the input 16 of the audio processing system 10 through the normal microphone channel 12 (which may be a wireless channel).

The signals S43 as emitted by the respective loudspeakers LS_L and LS_R are coded such as to make possible a distinction between the different loudspeakers. Thus, the audio processing system will receive a plurality of microphone pickup signals $S44_L$ and $S44_R$, arriving in a certain relative order and with a certain time difference. It is noted that the audio processing system 10 is also aware of the emission times, and is consequently capable of calculating traveling times. The audio processing system 10 is designed to filter out these microphone pickup signals $S44_L$ and $S44_R$ and process them in order to calculate the actual location of the detecting microphone 11 with respect to the respective locations of the loudspeakers LS_L and LS_R , on the basis of the respective traveling times of the microphone pickup signals $S44_L$ and $S44_R$.

The coding of the microphone bearing signals S43 may be any suitable coding, suitable for adequate distinction by the audio processing system 10. In one embodiment, the

microphone bearing signals S_{43L} and S_{43R} may be mutually identical but emitted at different times for different loudspeakers. In this respect, the repetition time between successive signals emitted from different loudspeakers may even be longer than the travelling and processing time. In another embodiment, the microphone bearing signals S_{43L} and S_{43R} may comprise pulse trains having mutually different carrier frequencies. In another embodiment, the microphone bearing signals S_{43L} and S_{43R} may comprise pulse trains containing a pulse width coding or a pulse distance coding.

A suitable coding is a coding of which the auto correlation function resembles a pulse as much as possible, such as a Barker code. It is also possible to use pairs of codes, such as for instance Golay codes; for more detailed information on Golay codes, reference is made to the article "The Merit Factor Of Long Low Autocorrelation Binary Sequences" by M.J.E. Golay in IEEE Transactions on Information Theory, (May 1982), vol.28, nr.3, p.543-549.

In any case, the audio processing system 10 is adapted to derive, from the signals S_{42} or S_{44} as received at its input 16, a signal or parameter value indicating the actual source location of person W. Herein, the implementation of said signal or value is not relevant: it may for instance be an analog value such as a voltage level, or alternatively it may be a digital value. In this respect, it is noted that it is not specifically necessary to actually calculate such location. It will be sufficient if the audio processing system 10 has access to a relationship between on the one hand the received microphone signals S_{42} or S_{44} and on the other hand an adequate balance ratio ρ_{LR} . This relationship may be stored in a memory 13 associated with the audio processing system 10, for instance in the form of a translation table. In order to allow for different arrangements of the loudspeakers, it is preferred that the relationship is adjustable, such that in a learning phase, for various microphone locations, adequate control settings can be determined and stored in said memory.

Evidently, apart from the pickup signals S_{44} , the microphone signal S_M also contains sound signals, indicated as voice signals SV , to be processed by the audio processing system 10 for reproduction through the loudspeakers LS_L and LS_R , for which purpose the audio processing system 10 generates loudspeaker drive signals SD_L and SD_R , respectively, at its outputs 17 and 18. Specifically, these loudspeaker drive signals SD_L and SD_R , respectively, comprise individual drive signals for the individual drivers of the loudspeakers. The audio processing system 10 is adapted to control the balance ratio ρ_{LR} of the loudspeaker drive signals SD_L and SD_R to the loudspeakers LS_L and LS_R such that a virtual spot location $AVSSL$ will be perceived corresponding to the actual source location of

person W as indicated by the received microphone signals S42 or S44. In the situation illustrated in Fig. 2, where the singer W is closer to the right-hand loudspeaker LS_R, the audio processing system 10 will decrease the balance ratio ρ_{LR} , either by a relative amplification of the loudspeaker drive signals SD_R to the right-hand loudspeaker LS_R, or by a relative attenuation of the loudspeaker drive signals SD_L to the left-hand loudspeaker LS_L, or both.

In the above, the present invention has been explained for the case of a listener L1, L2 in an audience, i.e. at a relatively large distance from the loudspeakers. For such listener, the invention provides in correlating the sound perception to the visual perception of viewing the person W. However, a similar problem exists with respect to the person W himself. This person W also is a listener in the sense that he will hear his own voice (or the music produced by his instrument) being reproduced by the loudspeaker systems. However, there are two important differences between the listener W and the listener L1, L2 in an audience. A first important difference is the fact that the listener W is relatively close the loudspeakers. A second important difference is the fact that the listener W has no visual perception of viewing a singer. Therefore, the speaker/singer W suffers from the strange perception of hearing his own voice (music) coming from outside himself, from a location not coinciding with his own location. Typically, the speaker/singer W will hear his own voice (music) as coming from a location (virtual source location) coinciding with the closest loudspeaker, unless he is located in the median plane M.

In some cases, it may be desirable that this problem is solved for the speaker/singer W instead of for the audience listeners L1, L2. It is noted that the technical measures proposed for providing a singer-related solution are different from the technical measures proposed for providing an audience-related solution, but it is noted that in both cases the technical measures are based on the same inventive concepts. Further, an audio processing system implementing the technical measures proposed for providing a singer-related solution may be physically distinct from an audio processing system implementing the technical measures proposed for providing an audience-related solution. However, in the preferred embodiment shown, the audio processing system 10 comprises a selection switch 14, and the audio processing system 10 is responsive to a signal received from this switch 14 to operate either in an "audience" mode as described above, or in a "singer" mode as will be discussed below.

Briefly stated, in the "singer" mode, the audio processing system 10 operates oppositely with respect to the "audience" mode. As explained in the above, in the "audience" mode the audio processing system 10 shifts the balance ratio ρ_{LR} towards the loudspeaker

(LS_R) closest to the speaker/singer W. In contrast, in the "singer" mode, the audio processing system shifts the balance ratio ρ_{LR} towards the loudspeaker (LS_L) which is the most remote from the speaker/singer W. This balance shift may be such that the speaker/singer W perceives a virtual source location independent of his actual location, such as for instance the center C between the two loudspeakers. Preferably however, the balance shift is such that the speaker/singer W perceives his own voice (music) as coming from his own location, i.e. an adaptive virtual singer source location AVSSL coinciding with his actual location.

In both cases, the audio processing system needs to be provided with location detection means in order to be able to calculate the required amount of balance shift. For these location detection means, and the way in which a relation between detection signals and balance shift is determined, the same applies as mentioned above with respect to the "audience" mode.

In an alternative preferred embodiment, the audio processing system is capable of solving the problem for the speaker/singer W as well as for the audience listeners 11, 12. In this preferred embodiment, schematically illustrated in Fig. 3A, the speaker/singer W is equipped with a headset 48 of earphones, and the audio processing system 10 has a further output 19 coupled to a transmitter 47 for transmitting, preferably wireless, an output signal to the headset 48. The selection switch 14 may be omitted in this case. With respect to its outputs 17 and 18 driving the loudspeakers LS_L and LS_R, the audio processing system 10 operates in the "audience" mode as described above. With respect to its output 19, the audio processing system 10 may operate in a mono mode, such that the speaker/singer W hears his own voice with equal timing and equal intensity from the left-hand earphone and from the right-hand earphone and will perceive his own voice as coming from his own location. In this case, too, the virtual source location as perceived by the speaker/singer W moves along with the actual source location, i.e. the actual location of the speaker/singer W.

It is noted that, in this embodiment, the transmitter 41 may be associated with the headset 48 instead of the microphone 11.

In the above, the present invention has been explained with respect to a single sound source, i.e. speaker/singer-generated signals only. This would apply, for instance, to a speaker addressing an audience, or a singer singing without accompanying music, or a single musician playing an instrument. The situation becomes more complicated in the case of two or more sound sources. Several situations are conceivable.

In a first situation, there is only one sound signal from a moving source (speaker/singer) and one or more signals from a stationary source, such as for instance in case

of a singer being accompanied by music. The stationary source may be a live orchestra, or a recording played from, for instance, CD. The recording may even comprise singing. In the following, such stationary audio will be called "background"; in contrast, the audio produced by the person W will be called "foreground".

In a second situation, there are two or more sound signals from moving sources, such as for instance in case of two or more individual singers moving on stage independently from each other. In this case, there may be stationary sources as well.

The present invention also provides solutions to these further complications.

Fig. 4 illustrates an embodiment of the audio processing system 10 comprising an input 51 for receiving background signals, for instance from a CD player. It is assumed that input 51 is a stereo input; therefore, two signal lines are shown. Although the audio processing system 10 is capable of receiving foreground signals from a stereo microphone, it is assumed here that microphone input 16 is a mono input; therefore, only one signal line is shown.

The recorded background may comprise audio of the same nature as the foreground. For instance, the person W may be a singer, and the recorded background may comprise singing. In another example, the person W may be a musician such as a violin player, and the recorded background may comprise violin music. In any case, the sound signals corresponding to this background audio having the same nature as the foreground will be indicated as "same nature background audio signals". If desired, these same nature background audio signals may be suppressed by a band reject filter 52. A suitable frequency range is, for example, 300 - 4500 Hz for the case of voice audio such as singing. Further, the microphone signals received at microphone input 16 may contain background, for instance because the microphone picks up the sound of an accompanying music band; if desired, these background signals may be suppressed by an echo feedback suppressor 53 or, alternatively, a band pass filter. In case the person W produces voice audio such as singing, a suitable frequency range is, for example, 300 - 4500 Hz.

In case the person W plays an instrument, other frequency ranges may be set in accordance with the typical frequency spectrum of such instrument, instead of the frequency ranges mentioned above, as should be clear to a person skilled in the art.

The audio processing system 10 comprises music processing means 54 for processing the background signals in a suitable manner; this background processing means 54 may be a conventional processing means, and is not discussed in great detail. The background processing means 54 has outputs 56L and 56R (stereo).

The audio processing system 10 comprises foreground processing means 55 for processing the microphone signals (voice; music) in a suitable manner; this foreground processing means 55 may be a conventional processing means, and is not discussed in great detail. The foreground processing means 55 may have two different outputs 57L and 57R for handling a stereo signal; however, in this embodiment, the foreground processing means 55 has one output 57.

The audio processing system 10 further comprises two controllable amplifiers 59L, 59R, both having their input connected to the output 57 of the foreground processing means 55. The audio processing system 10 comprises control means 60 generating control signals to the controllable amplifiers 59L, 59R such as to set the gains G_{59L} , G_{59R} of said controllable amplifiers 59L, 59R in response to the microphone bearing signals S42 or S44, the output signal from mode selection switch 14, and the information in memory 13, as will be clear to a person skilled in the art, such as to produce weighted foreground signals S_{57L} and S_{57R} . A first adder 61 has inputs connected to the left-hand output 56L of the background processing means 54 and to the output of left-hand amplifier 59L, and has its output connected to first output 17 to provide the left-hand loudspeaker drive signal SD_L . A second adder 62 has inputs connected to the right-hand output 56R of the background processing means 54 and to the output of right-hand amplifier 59R, and has its output connected to second output 18 to provide the right-hand loudspeaker drive signal SD_R .

The ratio G_{59L}/G_{59R} of said two gains corresponds to the balance ratio ρ_{LR} .

The embodiment shown in Fig. 4 is also capable of providing a headset drive signal SD_H to a headset output 19. A third adder 63 has inputs connected to the output 57 of the foreground processing means 55 and to the left-hand output 56L and the right-hand output 56R of the background processing means 54. All these signals are added without weighing. The mono signal at output 19 will generate a virtual source location between the earphones of the headset 48.

Thus, the audio processing system 10 illustrated in Fig. 4 comprises at least two signal processing paths for processing the movable microphone signals differently from the stationary source signals. Fig. 5 schematically illustrates a more complicated embodiment of audio processing system 10, capable of handling an arrangement of two or more movable microphones. Generally, Fig. 5 is comparable to Fig. 4, but the filters 52, 53 and processing units 54, 55 are omitted for sake of convenience. The audio processing system 10 has a plurality of microphone input ports 16i (in the example shown: three), and separate signal processing paths for each microphone signal, such signal processing path including

controllable amplifiers 59Li and 59Ri. For controlling the gains of the controllable amplifiers in each individual signal processing path, the audio processing system 10 will comprise corresponding control units 60i, responsive to a corresponding microphone location detection means (not shown). Thus, before being combined in adders 61 and 62, each microphone signal is individually processed as described earlier, in correspondence with the actual location of the corresponding microphone. In the overall signal as supplied to the loudspeakers LS_L and LS_R, each microphone signal component is reproduced by the loudspeakers LS_L and LS_R such that a corresponding adaptive virtual singer source location results, actually corresponding to the actual location of the corresponding singer.

In a preferred embodiment, the audio processing system 10 has a plurality of headset outputs 19i, each for supplying a headset output signal SDHi to a headset to be used by one respective singer. A processing unit 70 receives all weighted microphone signals S57Li and S57Ri. All headset output signals may be mono, and may be identical to each other. However, in a more sophisticated processing, the illusion of directivity regarding the other singers is created in each headset output signal. To that end, all headset output signals are preferably stereo. In respect of each headset output signal SDHi, the corresponding weighted microphone signals S57Li and S57Ri are added and supplied in mono. Also, possible background from background input 51 is supplied in mono. The microphone components of the other singers may be supplied with such left/right weighing that virtual source locations are perceived corresponding to the actual locations of the other singers.

It should be clear to a person skilled in the art that the present invention is not limited to the exemplary embodiments discussed above, but that various variations and modifications are possible within the protective scope of the invention as defined in the appending claims.

CLAIMS:

1. Sound producing system, comprising:
 - a microphone;
 - an audio processing system having a microphone input for receiving a microphone signal from the microphone, the audio processing system further having loudspeaker outputs for outputting loudspeaker drive signals reproducing at least part of the microphone signal from the microphone;
 - wherein said audio processing system is adapted to generate said loudspeaker drive signals with a certain adaptable balance ratio;
 - the sound producing system further comprising source location detection means associated with the audio processing system for generating a location signal representing the actual source location of the microphone;
 - wherein the audio processing system is responsive to the location signal received from said source location detection means to amend the setting of said balance ratio.

2. Sound producing system according to claim 1, wherein the audio processing system is associated with a memory containing information regarding a relationship between on the one hand said location signal and on the other hand an adequate balance ratio.

3. Sound producing system according to claim 2, wherein said relationship is adjustable.

4. Sound producing system according to any of claims 1-3, wherein said actual source location detection means comprise:
 - a mobile transmitter to be carried by a microphone user, capable of emitting a predefined signal;
 - and at least one receiver coupled to the audio processing system, each receiver being capable of receiving said predefined signal and sending receiver signals to the audio processing system;

wherein the audio processing system is adapted to derive a location information from said receiver signals, for instance from the relative order and time differences of said receiver signals obtained from a plurality of receivers.

5. Sound producing system according to claim 4, wherein said transmitter is associated with the microphone.
6. Sound producing system according to claim 4, wherein said transmitter is associated with a headset.
7. Sound producing system according to any of claims 1-3, further comprising a set of loudspeakers, connected to said loudspeaker outputs, respectively, adapted to generate sound in response to said loudspeaker drive signal;
wherein said audio processing system is designed to generate predetermined microphone bearing signals through the loudspeakers;
and wherein said audio processing system is designed to derive a location signal from the relative order and time differences of microphone pickup signals, corresponding to the microphone bearing signals as emitted by the respective loudspeakers.
8. Sound producing system according to claim 7, wherein said location bearing signals are generated within a frequency range inaudible to the human ear.
9. Sound producing system according to claim 7 or 8, wherein the audio processing system is designed to take into account the emission time points of the microphone bearing signals, and is capable of calculating traveling times.
10. Sound producing system according to any of claims 1-9, operable in an audience mode, wherein the audio processing system is adapted to shift said balance ratio towards a loudspeaker closest to the actual source location of the microphone.
11. Sound producing system according to claim 10, wherein the audio processing system is adapted to amend the setting of said balance ratio such that an adaptive virtual source location substantially corresponds with the actual source location of the microphone.

12. Sound producing system according to any of claims 1-9, operable in a singer mode, wherein the audio processing system is adapted to shift said balance ratio towards a loudspeaker system most remote from the actual source location of the microphone.

13. Sound producing system according to claim 12, wherein the audio processing system is adapted to amend the setting of said balance ratio such that an adaptive virtual spot location, as perceived by a user of the microphone, substantially coincides with the actual source location of the microphone.

14. Sound producing system according to claim 12, wherein the audio processing system is adapted to amend the setting of said balance ratio such that a virtual spot location, as perceived by a user of the microphone, substantially coincides with a substantially constant location independent of the actual microphone location, such as for instance the center between the two loudspeakers.

15. Sound producing system according to any of the previous claims 1-14, further comprising a mode selection switch, the audio processing system being responsive to a signal received from said switch to operate either in an "audience" mode or in a "singer" mode.

16. Sound producing system according to any of the previous claims 1-14, wherein said audio processing system further comprises a headset output for outputting headset drive output signals reproducing at least part of the microphone signal from the microphone, and wherein the audio processing system is adapted to generate said headset drive output signals with a constant setting of said balance ratio, or to supply mono output signals to said headset output, such that a virtual spot location, as perceived by a user of the headset, coincides substantially with the center between the headset earphones.

17. Sound producing system according to any of the previous claims 1-16, comprising at least one microphone input and at least one background input for receiving background signals;

the audio processing system having a first signal processing path for processing any background signals received at said at least one background input, and for generating loudspeaker drive signals corresponding to said background signals at a constant first balance ratio;

the audio processing system having a second signal processing path for processing any microphone signals received at said at least one microphone input, and for generating loudspeaker drive signals corresponding to said microphone signals at an adaptable second balance ratio differing from the first balance ratio;

wherein the audio processing system is responsive to the location signal received from the source location detection means to amend the setting of said second balance ratio.

18. Sound producing system according to claim 17, further comprising a band reject filter associated with the background input for suppressing foreground signals, wherein the band reject filter has a reject band corresponding to the frequency range of said foreground signals, for instance in the order of 300 - 4500 Hz.

19. Sound producing system according to claim 17 or 18, further comprising an echo suppressor or a band pass filter associated with the microphone input for suppressing background signals, wherein the suppressor or band pass filter preferably has a pass band corresponding to the frequency range of foreground signals, for instance in the order of 300 - 4500 Hz.

20. Sound producing system according to any of the previous claims 1-19, comprising a plurality of microphone inputs and at least two source location detection means corresponding to respective microphones;

the audio processing system having a plurality of signal processing paths for processing respective microphone signals received at a respective one of said microphone inputs;

the audio processing system being adapted to generate loudspeaker drive signals for said loudspeakers, the loudspeaker drive signals comprising a plurality of loudspeaker drive signal components, each loudspeaker drive signal component corresponding to a respective one of said microphone signals;

wherein said audio processing system is adapted to generate said loudspeaker drive signal components at respective adaptable balance ratios;

the sound producing system further comprising source location detection means associated with the audio processing system for generating respective location signals representing the actual source locations of the respective microphones;

wherein the audio processing system is responsive to said location signals received from said source location detection means to amend the settings of said balance ratios.

21. Sound producing system according to claim 20, wherein the audio processing system is adapted to amend the settings of said balance ratios such that respective adaptive virtual source locations substantially correspond with the respective actual source locations of the respective microphones.

1/5

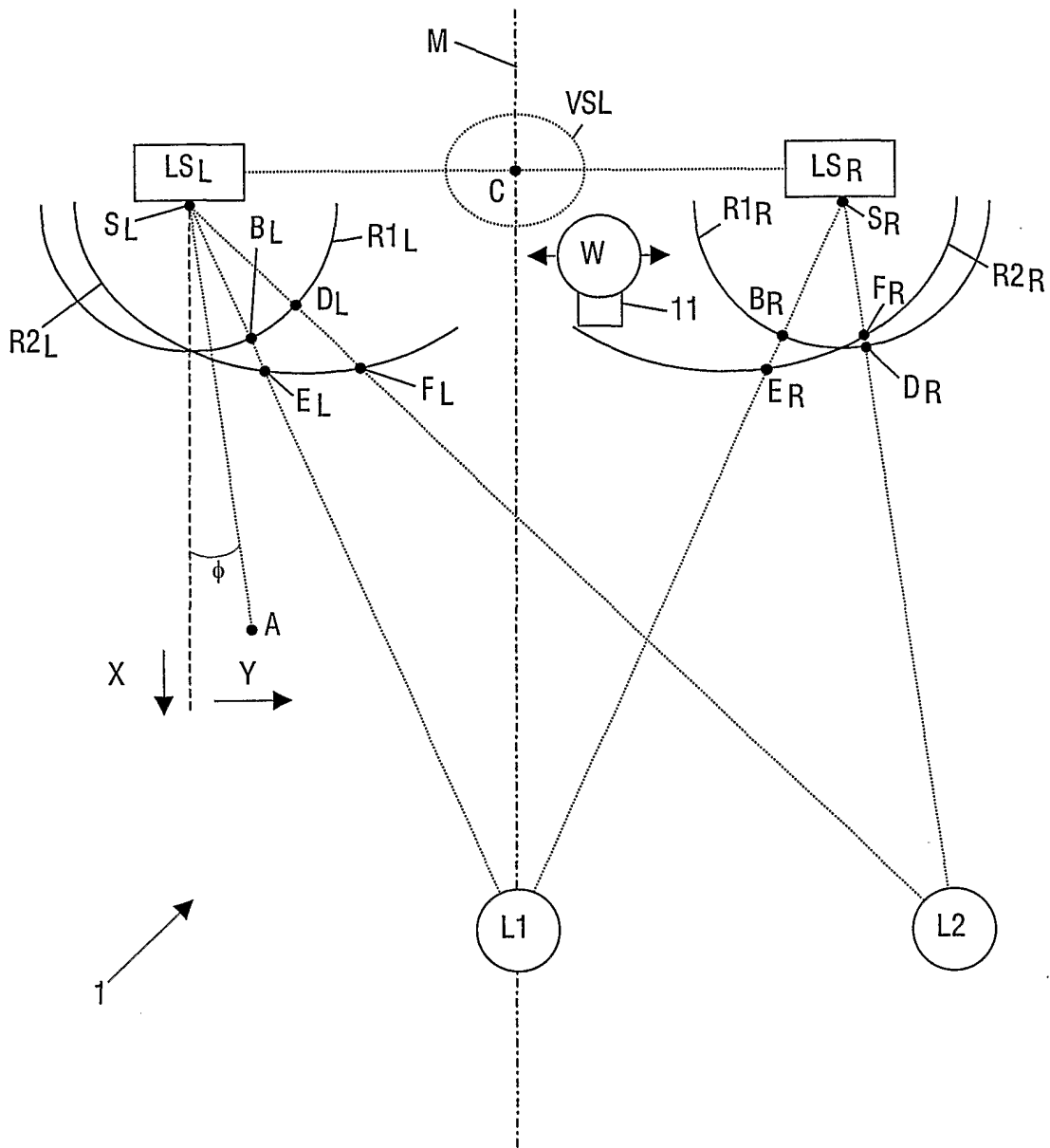


FIG.1

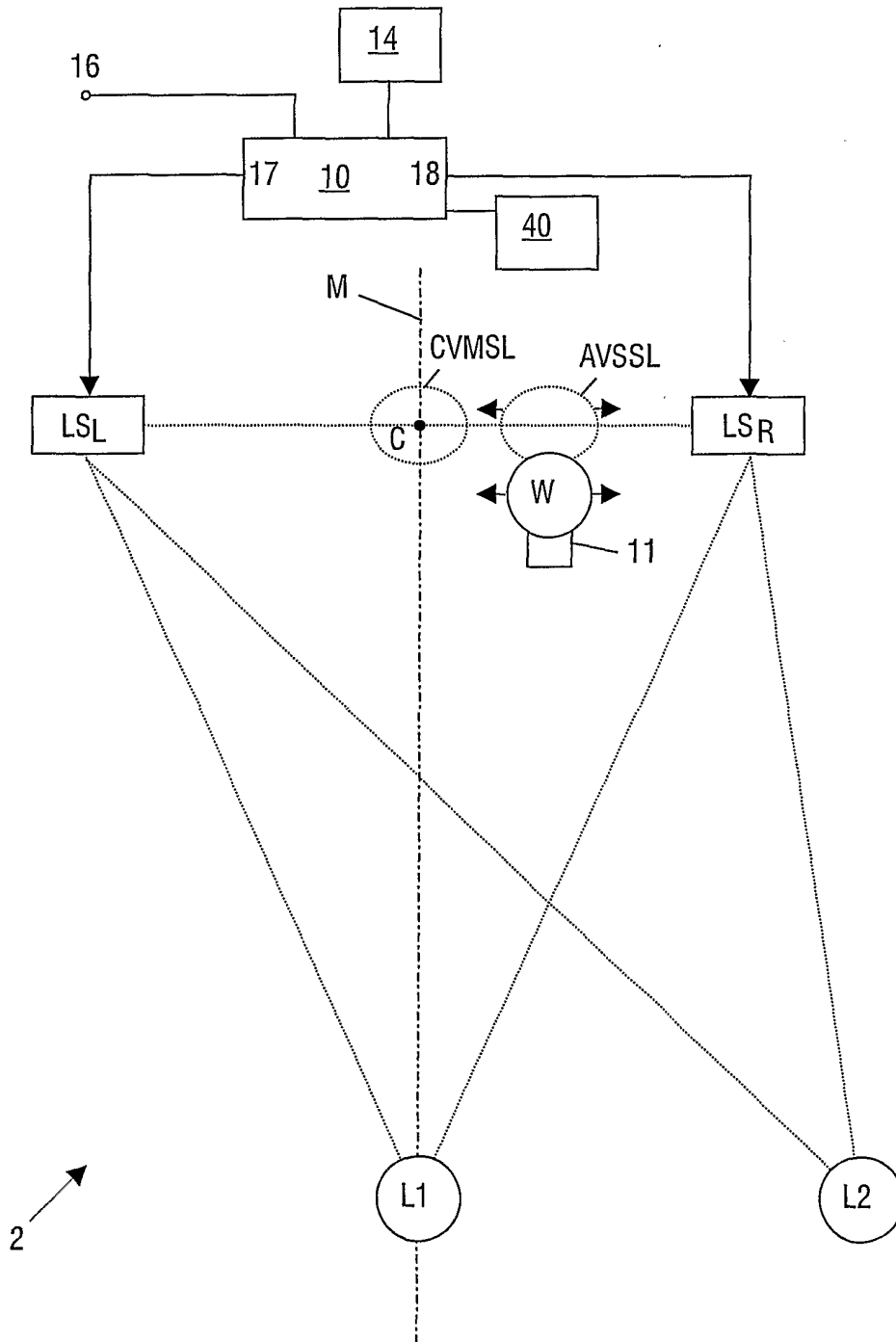


FIG.2

3/5

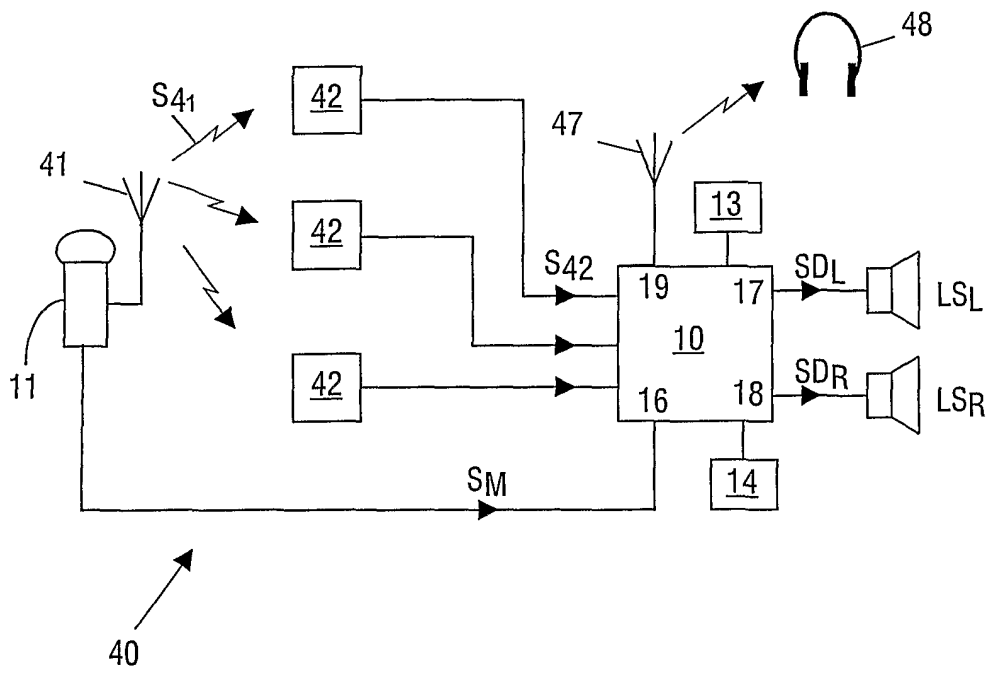


FIG.3A

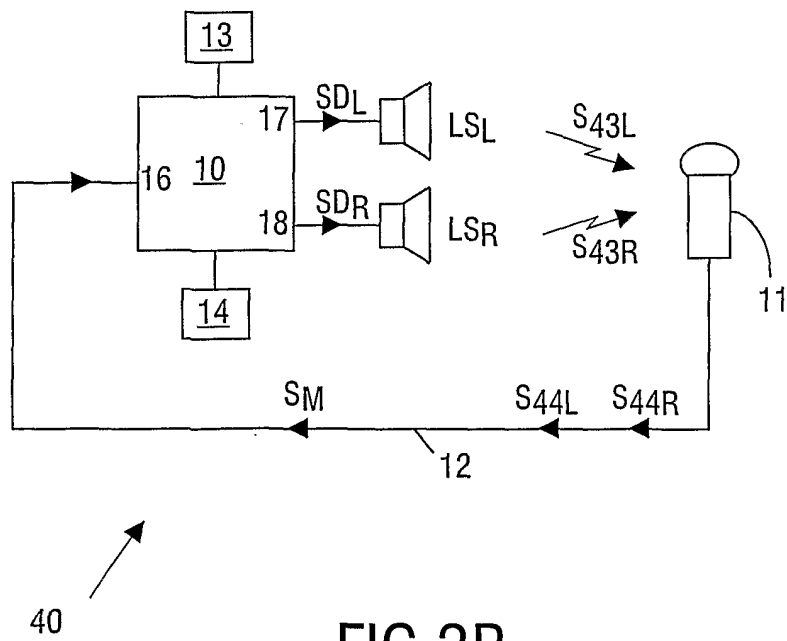


FIG.3B

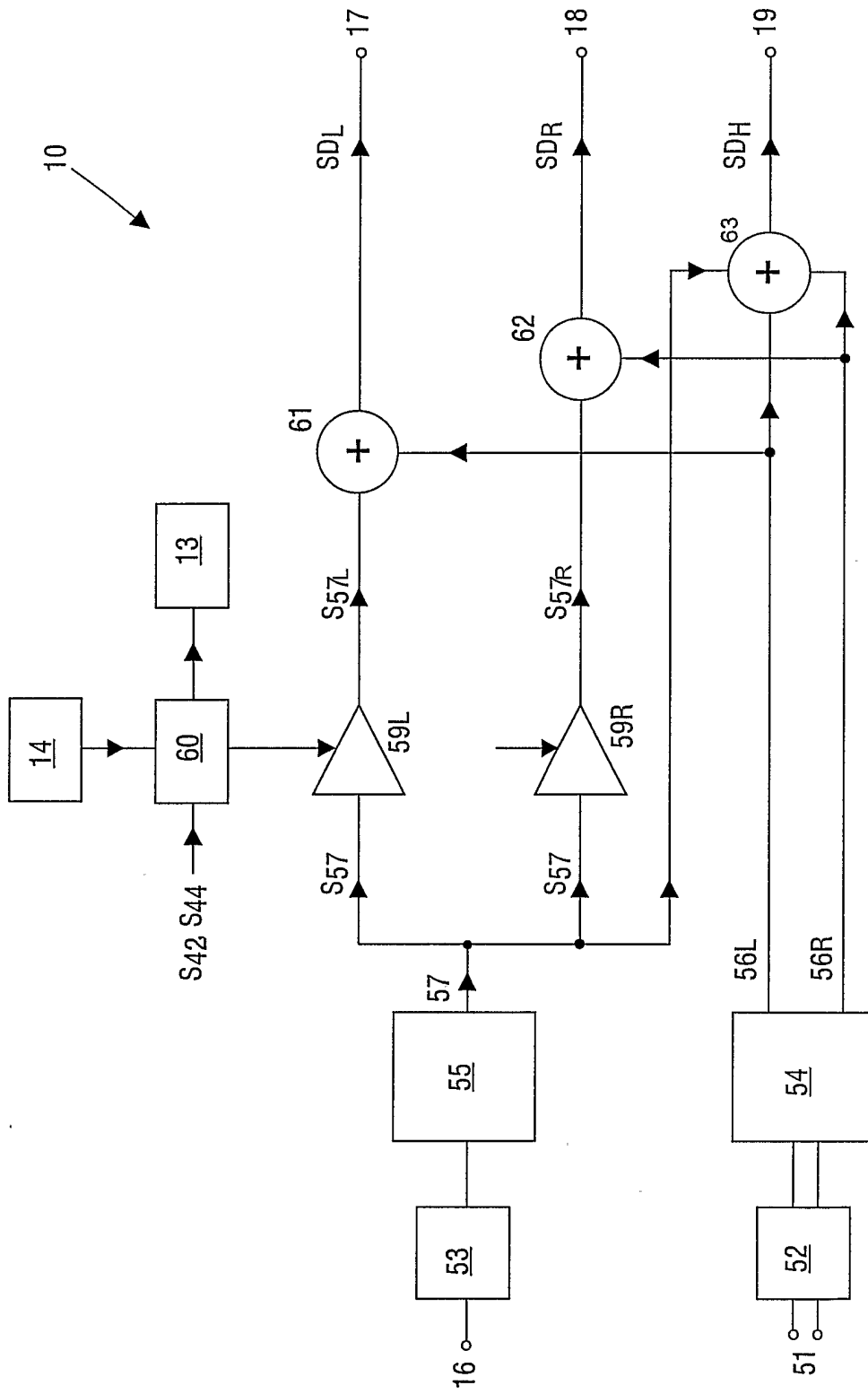


FIG. 4

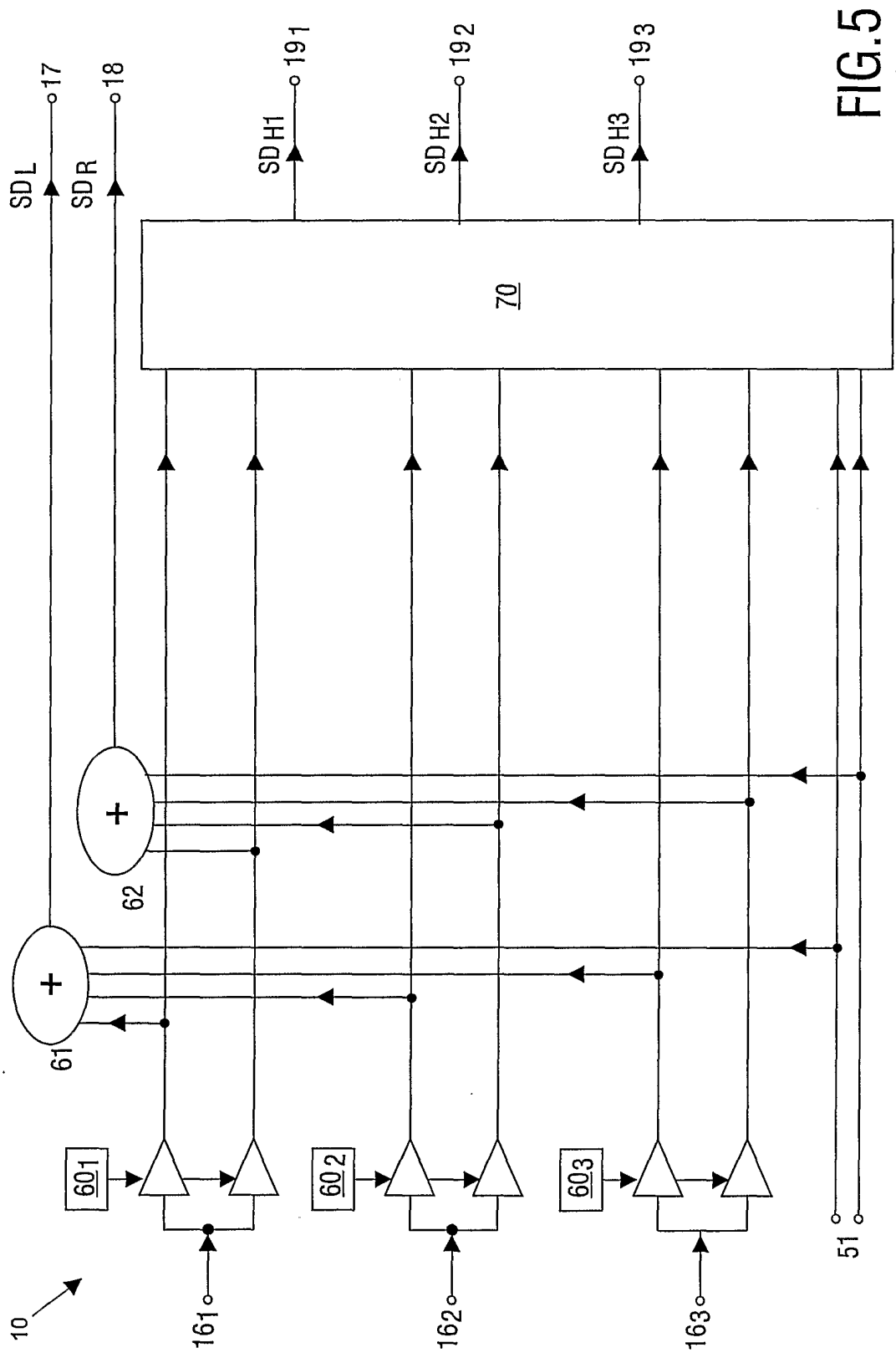


FIG. 5

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/IB/03/03389

A. CLASSIFICATION OF SUBJECT MATTER
 IPC 7 H04S1/00 G10H1/36

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
 IPC 7 H04S G10H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

PAJ, EPO-Internal, WPI Data, INSPEC, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	PATENT ABSTRACTS OF JAPAN vol. 018, no. 327 (P-1757), 21 June 1994 (1994-06-21) -& JP 06 075585 A (KENWOOD CORP), 18 March 1994 (1994-03-18) the whole document	1-3, 7, 8, 12, 13
X	PATENT ABSTRACTS OF JAPAN vol. 2000, no. 05, 14 September 2000 (2000-09-14) & JP 2000 039894 A (YAMAHA CORP), 8 February 2000 (2000-02-08)	14
A	abstract	12, 16, 17

Further documents are listed in the continuation of box C:

Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

& document member of the same patent family

Date of the actual completion of the international search

23 October 2003

Date of mailing of the international search report

03/11/2003

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
 NL - 2280 HV Rijswijk
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
 Fax: (+31-70) 340-3016

Authorized officer

Gerken, S

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IB/03/03389

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 386 478 A (PLUNKETT BRADLEY J) 31 January 1995 (1995-01-31) abstract column 2, line 23 -column 4, line 16; figure 1	9
A	----- PATENT ABSTRACTS OF JAPAN vol. 018, no. 513 (P-1805), 27 September 1994 (1994-09-27) & JP 06 175673 A (KENWOOD CORP), 24 June 1994 (1994-06-24) abstract	1,10
A	----- PATENT ABSTRACTS OF JAPAN vol. 016, no. 138 (E-1186), 7 April 1992 (1992-04-07) & JP 03 297300 A (PIONEER ELECTRON CORP), 27 December 1991 (1991-12-27) abstract	17,18
A	----- PATENT ABSTRACTS OF JAPAN vol. 1998, no. 03, 27 February 1998 (1998-02-27) & JP 09 297589 A (EKUSHINGU:KK;BROTHER IND LTD), 18 November 1997 (1997-11-18) abstract	20

INTERNATIONAL SEARCH REPORT

International Application No.

PCT/IB/03/03389

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
JP 06075585	A	18-03-1994	JP 2958930 B2	06-10-1999
JP 2000039894	A	08-02-2000	CN 1258905 A , B	05-07-2000
			KR 2000011932 A	25-02-2000
			US 6127618 A	03-10-2000
US 5386478	A	31-01-1995	NONE	
JP 06175673 4	A		NONE	
JP 03297300 4	A		NONE	
JP 09297589 4	A		NONE	