

(19)



(11)

EP 1 538 867 B1

(12)

EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention
of the grant of the patent:
18.07.2012 Bulletin 2012/29

(51) Int Cl.:
H04R 3/00 (2006.01) **H04R 1/40** (2006.01)

(21) Application number: **03022273.1**

(22) Date of filing: **01.10.2003**

(54) **Handsfree system for use in a vehicle**

Freisprechanlage für ein Kraftfahrzeug

Système mains libres utilisé dans un véhicule

(84) Designated Contracting States:
**AT BE BG CH CY CZ DE DK EE ES FI FR GB GR
HU IE IT LI LU MC NL PT RO SE SI SK TR**

(30) Priority: **30.06.2003 EP 03014846**

(43) Date of publication of application:
08.06.2005 Bulletin 2005/23

(73) Proprietor: **Nuance Communications, Inc.
Burlington, MA 01803-4613 (US)**

(72) Inventor: **Christoph, Markus
94315 Straubing (DE)**

(74) Representative: **Grünecker, Kinkeldey,
Stockmair & Schwanhäusser
Leopoldstrasse 4
80802 München (DE)**

(56) References cited:
**WO-A-01/84881 WO-A-03/015464
US-A- 5 715 319 US-A- 5 835 607**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 1 538 867 B1

Description

[0001] The invention is directed to a handsfree system for use in a vehicle comprising a microphone array with at least two microphones and a signal processing means.

[0002] For making telephone calls in a car, handsfree systems are used more and more since they provide increased comfort and reduce the risk of an accident as the driver is distracted only marginally. Because of that, in many countries, handsfree devices are even required by law.

[0003] WO 03/015464 discloses such a handsfree system for use in a vehicle.

[0004] Usually, a handsfree system comprises a microphone that can be fastened to a user such as the driver.

[0005] Due to the relatively large distance between the speaker's mouth and the microphone, many handsfree devices today suffer from the drawback of a poor speech quality. This is particularly due to the fact that in a car, usually a large ambient noise is present interfering with the speech signal. The noise stems from different sources such as the motor, wind, or car radio.

[0006] However, common methods for noise reduction are often costly to implement and require a large amount of memory and computing power. In particular, a signal processed by conventional noise reduction systems has a relatively large delay time which makes these systems unsuitable for real time applications, i.e. telephone applications.

[0007] It is, therefore, the problem underlying the invention to overcome the above drawbacks and provide a handsfree system for use in a vehicle with improved speech quality.

[0008] This problem is solved by a handsfree system according to claim 1.

[0009] In the context of this invention, the term "connected" also includes the case that a filter or another signal processing means is provided along the signal path between two devices or means. A beamformer processes signals emanating from a microphone array to obtain a combined signal. A beamformer comprises a beamsteering means being responsible for time delay compensation of the different microphones and a summing means. In its simplest form (Delay-and-Sum beamformer), beamforming only comprises delay compensation and summing of the compensated signals. Beamforming allows to provide a specific directivity pattern for a microphone array. Usually, a beamformer can be implemented as digital system with a plurality of digital filter using, for example, digital signal processors (DSP). A beamformer can be configured as an adaptive or a non-adaptive beamformer. Adaptive means that relevant parameters such as filter coefficients can be re-calculated during use of the system in order to adapt the beamformer to changing conditions. In the non-adaptive case, the system parameters are determined once by calibrating the beamformer and, then, kept unchanged. In both cases of a non-adaptive and an adaptive beamformer, the beamforming, in principle, can be performed in the time domain or in the frequency domain.

[0010] A handsfree system in accordance with the invention shows an excellent acoustic performance in a vehicular environment. Due to the beamformer, an improved directivity is obtained and, furthermore, speech signals are enhanced and ambient noise is reduced. The adaptive post-filter (responsible for filtering a signal after the beamforming) further reduces the noise in the signal.

[0011] According to a preferred embodiment, the adaptive post-filter can be a filter in the time domain. If the post-filtering is performed in the time domain, the delay time is reduced and the implementation is simplified.

[0012] The adaptive post-filter is a Wiener filter. It turns out that a Wiener filter is particularly suitable for filtering in a car environment.

[0013] In order to reduce spectral distortions of the filtered signal, the adaptive post-filter is a linear-phase filter. In particular, the adaptive post-filter is a linear-phase Wiener filter.

[0014] According to a preferred embodiment, the signal processing means can further comprise at least two adaptive filters having an input connected to the output of the beamsteering means and an output connected to the adaptive post-filter, wherein the at least two adaptive filters are configured to determine adaptive filter parameters for the adaptive post-filter.

[0015] In this way, background filters are provided for adaptively estimating the filter parameters for the adaptive post-filter.

[0016] Preferably, for each of the at least two microphones, an adaptive filter can be provided having an input connected to the output of the beamsteering means. Thus, for each output of the beamsteering signal corresponding to a microphone, adaptive filter parameters can be determined for the adaptive post-filter. The actual filter parameters of the post-filter can be given, for example, by the filter parameters determined by one of the adaptive filters or the mean of the filter parameters determined by several different adaptive filters.

[0017] Advantageously, an input of each of the at least two adaptive filters can be further connected to the output of the beamformer. This allows for an adaption of the respective filter parameters directly with respect to the beamformed signal.

[0018] According to a preferred embodiment, the signal processing means can further comprise a pre-emphasis filter, in particular, comprising a pre-whitening filter, having an input connected to an output of the adaptive post-filter and/or a pre-emphasis filter, in particular, comprising a pre-whitening filter, having an input connected to the output of the

beamsteering means and an output connected to the at least two adaptive filters.

[0019] Such a pre-emphasis filter, on the one hand, emphasizes high frequencies and, on the other hand, attenuates low frequencies which is particularly useful to reduce low frequency correlated noise. Preferably, the pre-emphasis filter can comprise a pre-whitening filter. A pre-whitening filter whitens the spectral distribution of a signal. The filter coefficients of such a pre-whitening filter can be determined using a linear predictive coding (LPC) analysis, for example, via an adaptive lattice predictor (ALP) algorithm.

[0020] According to a preferred embodiment of the above handsfree systems, the signal processing means can further comprise an inverse filter, particularly a warped inverse filter. These filters are especially useful to adjust the microphone transfer function and to match the microphones of the array in this way. Preferably, the beamformer can comprise at least one inverse filter, in particular, having an output for providing an inversely filtered signal to a summing means.

[0021] In order to overcome the matching problem, alternatively or additionally, matched microphones on the basis of silicone or paired microphones may be used.

[0022] The susceptibility of microphone arrays often increases with decreasing frequency. Due to this, a higher matching precision is preferred for low frequencies compared to high frequencies. A frequency depending adjustment of the microphone transfer functions with the use of warped filters reduces the required memory compared to the case of conventional FIR filters.

[0023] Preferably, each inverse filter can be an approximate inverse of a non-minimum phase filter. This results in an inverse filter which is both stable and has no phase error.

[0024] According to a preferred embodiment, an inverse filter may be combined with another filter of the handsfree system, for example, a filter of the beamformer. Such a combination in one filter results in a simplified implementation.

[0025] Preferably, the signal processing means of the above handsfree systems can comprise a non-adaptive post-filter having an input connected to an output of the adaptive post-filter. The non-adaptive post-filter may directly follow the adaptive post-filter. Such a filter is used to compensate for the ambient acoustics of a speaker. Thus, the non-adaptive post-filter may have the form of an inverse room filter.

[0026] In order to further reduce low frequency noise, according to a preferred embodiment, the signal processing means may further comprise an adaptive noise canceller (ANC), for electrical ANC implementations.

[0027] Preferably, the ANC can be connected to a non-acoustic sensor to determine a noise signal, for example, by using the tachometer of the vehicle. The ANC, advantageously, can have an output connected to the input of the beamformer and/or of the adaptive post-filter.

[0028] For a further improvement of the speech signal quality, the signal processing means of the previously described handsfree systems can comprise an acoustic echo canceller AEC. Preferably, the AEC can comprise an echo shaping filter. In this way, a frequency selected echo attenuation may be obtained. As in the case of an ANC, the AEC can have an output connected to the input of the beamformer and/or of the adaptive post-filter.

[0029] According to a preferred embodiment of all previously described handsfree systems, the beamformer can be a non-adaptive beamformer. By using a non-adaptive beamformer with fixed filters, the computing power during operation of the system is reduced.

[0030] Preferably, the beamformer may be a superdirective beamformer which further improves the acoustic performance.

[0031] Advantageously, the beamformer may be a regularized superdirective beamformer using a finite regularization parameter μ . The regularization parameter usually enters the equation for computing the filter coefficients or, alternatively, is inserted into the cross-power spectrum matrix or the coherence matrix. In contrast to the maximum superdirective beamformer ($\mu = 0$), the regularized superdirective beamformer has reduced noise and is less sensitive to an imperfect matching of the microphones.

[0032] The finite regularization parameter μ , preferably, may depend on the frequency. This achieves an improved gain of the array compared to a regularized superdirective beamformer with fixed regularization parameter μ . According to a preferred embodiment, each superdirective filter may result from an iterative design based on a predetermined maximum susceptibility. This enables an optimal adjustment of the microphones, particularly with respect to the transfer function and the position of each microphone.

[0033] By using a predetermined maximum susceptibility, defective parameters of the microphone array can be taken into account to further improve the gain. The maximum susceptibility may be determined as a function of the error in the transfer characteristic of the microphones, the error in the microphone positions and a predetermined (required) maximum deviation in the directional diagram of the microphone array. The time-invariant impulse response of the filters will be determined iteratively only once; there is no adaption of the filter coefficients during operation.

[0034] According to a preferred embodiment, each superdirective filter can be a filter in the time domain. Filtering in the frequency domain is a possible alternative, however, requiring to perform a Fourier transform (FFT) and an inverse Fourier transform (IFFT), thus, increasing the required memory.

[0035] Advantageously, the beamformer may have the structure of a generalized sidelobe canceller (GSC). In this way, at least one filter can be saved. The implementation in the GSC structure, however, is only possible in the frequency

domain.

[0036] In order to obtain an optimal adaption of the handsfree system to a particular noise situation, according to a preferred embodiment, the beamformer can be a minimum variance distortionless response (MVDR) beamformer.

[0037] According to a preferred embodiment, the microphone array can comprise at least two microphones being arranged in endfire orientation with respect to a first position. An array in endfire orientation has a better directivity and is less sensitive to a mismatched propagation or delay time compensation. The first position can be the location of the drivers head, for example.

[0038] Preferably, the microphone array can comprise at least two microphones being arranged in endfire orientation with respect to a second position. Thus, the handsfree system of the invention has a good directivity in two directions. Speech signals coming from two different positions, for example, from the driver and the front seat passenger, can both be recorded in good quality.

[0039] According to a preferred embodiment, the signal processing means may comprise at least two beamformers. A first beamformer may be used for signals from a first position and a second beamformer may be used for signals from a second position. In this case, advantageously, the handsfree system may further comprise a voice activity detector (VAD) and/or a switch control means. The switch control and the VAD are used to determine how to combine the output of the at least two beamformers.

[0040] Advantageously, the handsfree system can comprise a residual echo suppression (RES) means and/or a dynamic volume control (DVC). A RES means serves for suppression of residual echoes, in particular, being present in the signal resulting from the adaptive post filter. Thus, a residual echo suppression means can comprise an input connected to the output of the adaptive post filter. Furthermore, a RES means can comprise an input for receiving a far end signal. A DVC is intended for dynamically adapting the output volume of a far end signal depending on the ambient noise level being present in the vehicle.

[0041] According to a preferred embodiment, the at least two microphones in the first endfire orientation (endfire orientation with respect to a first position) and the at least two microphones in the second endfire orientation (endfire orientation with respect to a second position) can have a microphone in common. In this way, already a microphone array consisting of only three microphones provides an excellent directivity for use in a vehicular environment.

[0042] According to a preferred embodiment of all previously discussed handsfree systems, the microphone array may comprise at least two subarrays. Each subarray of microphones may be optimized for a specific frequency band yielding an improved overall directivity.

[0043] To decrease the total number of microphones, preferably, at least two subarrays may have at least one microphone in common.

[0044] According to a preferred embodiment, the above handsfree systems may comprise a frame wherein each microphone of the microphone array is arranged in a predetermined, in particular fixed, position in or on the frame. This ensures that after manufacture of the frame with the microphone, the relative positions of the microphones are known. Such an array can be easily mounted in a vehicular cabin.

[0045] According to a preferred embodiment, at least one microphone may be a directional microphone. The use of directional microphones improves the array gain.

[0046] Preferably, at least one directional microphone may have a cardioid characteristic. This further improves the array gain. More preferred, the cardioid characteristic is a hyper-cardioid characteristic.

[0047] Advantageously, at least one directional microphone may be a differential microphone. This results in a microphone array with excellent directivity and small dimensions, in particular, the differential microphone may be a first order differential microphone.

[0048] The invention is further directed to a vehicle, particularly a car, comprising any of the above-described handsfree systems.

[0049] The invention is also directed to the use of any of the previously described handsfree systems in a vehicle, in particular, a car.

[0050] Furthermore, the invention provides a method for noise reduction in a vehicular handsfree system according to claim 21.

[0051] This method results in an excellent acoustic performance of a handsfree system in a vehicular environment.

[0052] According to a preferred embodiment, the adaptive filtering can be performed in the time domain. In this way, particularly the delay time is reduced.

[0053] Preferably, the method can further comprise

providing at least two adaptive filters, particularly Wiener filters, wherein beam processing the input signals by a beamformer forming comprises beamsteering the input signals for providing beamsteered signals corresponding to one of the at least two microphones and summing the signals, and adaptively filtering comprises receiving and processing at least one beamsteered signal by at least one of the at least two adaptive filters to determine adaptive filter parameters for the adaptive post filter.

[0054] According to a preferred embodiment, adaptively filtering can further comprise receiving a signal resulting from the beamformed signal by at least one adaptive filter and wherein processing the beamsteered signal can comprise determining adaptive filter parameters using the at least one beamsteered signal and the signal resulting from the beamformed signal.

[0055] Preferably, for each beamsteered signal, an adaptive filter can be provided for determining adaptive filter parameters using the beamsteered signal and the signal resulting from the beamformed signal.

[0056] In order to reduce low frequency correlated noise, receiving at least one beamsteered signal by at least one of the at least two adaptive filters can comprise processing the at least one beamsteered signal by a pre-emphasis filter, in particular, comprising a pre-whitening filter.

[0057] According to an advantageous embodiment, the above methods can further comprise processing a signal resulting from the microphone array by an inverse filter, in particular, a warped inverse filter.

[0058] Preferably, the methods can further comprise non-adaptively filtering a signal resulting from the adaptively filtered signal and/or processing a signal resulting from the adaptively filtered signal by a pre-emphasis filter.

[0059] The above method, advantageously, can further comprise processing a signal resulting from the microphone array, particularly resulting from the beamformed signal, by an adaptive noise canceller (ANC) and/or an acoustic echo canceller (AEC) and/or a residual echo suppression (RES) means.

[0060] According to a preferred embodiment, the input signals can be processed by a non-adaptive and/or superdirective and/or minimum variance distortionless response (MVDR) beamformer.

[0061] The invention also provides a computer program product comprising one or more computer readable media having computer-executable instructions for performing the steps of the above described methods.

[0062] Additional features and advantages will be described with reference to the examples illustrated in the drawings:

Fig. 1 illustrates the structure of a handsfree system according to the invention with an adaptive post-filter in the time domain;

Fig. 2 shows the structure of a beamformer in the frequency domain;

Fig. 3 illustrates an FXLMS algorithm;

Fig. 4 shows the structure of a beamformer in the time domain;

Figs. 5A, 5B illustrate preferred embodiments of arrangements of the microphone array in a vehicle;

Figs. 6A, 6B illustrate preferred embodiments of arrangements of a microphone array in a mirror;

Fig. 7 shows a microphone array consisting of three subarray;

Fig. 8 illustrates a superdirective beamformer in a GSC structure;

Fig. 9 illustrates a microphone array with two microphones in a noise field with a noise free sector;

Fig. 10 shows the structure of a superdirective beamformer comprising four first order gradient microphones;

Fig. 11 illustrates the structure of a handsfree system with an electrical ANC;

Fig. 12 shows the structure of an ANC;

Fig. 13 shows the structure of an embodiment of a handsfree system according to the invention with an ANC and AEC;

Fig. 14 illustrates the structure of an AEC; and

Fig. 15 shows another embodiment of a handsfree system according to the invention.

[0063] An example of the handsfree system in accordance with the present invention is shown in Fig. 1. In the following, first, the general structure will be shortly described, and, then, the different components will be explained in more detail. In the figures, it is to be noted that the dotted lines encasing some elements simply serve for better understanding of the figures without necessarily implying any actual combination or separation of different elements.

[0064] The main components of the system are a microphone array, a beamformer and an adaptive post-filter in the time domain. The microphone array 101, in this example, comprises four microphones 102. Each microphone 102 yields an output signal $x_i[k]$. The microphone signals may be filtered by an optional high-pass filter 103.

[0065] Then, the signals are passed to a beamformer. This beamformer may be a conventional delay and sum beamformer. However, in the present example, a preferred superdirective beamformer is shown. Such a beamformer comprises beamsteering means 104 and filters 105. The output signals of the beamformer may be passed through optional inverse filters 106 and, then, are summed by summing means 107 to yield a resulting beamformed signal $x[k]$.

[0066] This signal is passed through an adaptive post-filter 108 in the time domain which may be followed by an optional non-adaptive post-filter 109 and/or by an optional pre-emphasis filter (not shown). The adaption of the post-filter 108 is performed using a set of Wiener filters 110. The input signals of the Wiener filters 110 comprise, on the one hand, the individual signals resulting from the different microphones and, on the other hand, the summed signal $x[k]$. In the present example, the microphone signals are taken after the beamsteering. However, if the beamformer comprises further (superdirective) filters 105 as in the present case, it is also possible to take the microphone signals after this additional filtering. Before being presented to the Wiener filters, the microphone signals are passed through an optional pre-emphasis filter 111.

[0067] In the following, the functioning of a Wiener filter will be explained. A microphone signal $x[k]$ is the sum of the speech signal $s[k]$ and the noise $n[k]$. The microphone signal will be filtered by an impulse response $w(i)$ to obtain a noise reduced signal $\tilde{s}[k]$. It is the aim to minimize the mean square error between the undisturbed speech signal $s[k]$ and the output signal $\tilde{s}[k]$:

$$E\{e^2[k]\} = E\{s[k] - \tilde{s}[k]\}^2 = E\left\{\left(s[k] - \sum_{i=-\infty}^{\infty} w(i)x[k-i]\right)^2\right\} = \min.$$

[0068] In other words, the partial derivative of the mean square error with respect to the coefficients of the impulse response has to vanish yielding the Wiener-Hopf equation:

$$\sum_{i=-\infty}^{\infty} w(i)r_{xx}(l-i) = r_{sx}(l),$$

wherein $r_{xx}(l)$ and $r_{sx}(l)$ are the auto-correlation function and the cross-correlation function of the microphone signal and the undisturbed speech signal. One may assume that the speech signal and the noise are statistically independent, i.e. $r_{sx}(l) = r_{ss}(l)$, thus,

$$\sum_{i=-\infty}^{\infty} w(i)r_{xx}(l-i) = r_{ss}(l).$$

[0069] A transformation of this equation into the frequency domain yields the frequency response of the Wiener filter:

$$W(\omega) = \frac{\Phi_{sx}(\omega)}{\Phi_{xx}(\omega)} = \frac{\Phi_{ss}(\omega)}{\Phi_{sx}(\omega) + \Phi_{nn}(\omega)}.$$

[0070] In order to obtain a time variant filter, the power spectral densities in the above equation may be replaced by the corresponding short-time estimated values that may be obtained, for example, by a recursive averaging:

$$W(\kappa, \nu) = \frac{\overline{E}\{S(\kappa, \nu)^2\}}{\overline{E}\{X(\kappa, \nu)^2\}}.$$

wherein $S(\kappa, \nu)$ and $X(\kappa, \nu)$ are short-time spectra that may be determined, for example, with the help of DFT filter banks or an FFT. Here, κ is the time index and ν the frequency index; $\overline{E}\{\cdot\}$ represents the short-time average that may be obtained, for example, with the help of a first order IIR filter.

[0071] The short-time auto power spectral density of the speech signal in the numerator of the above equation is to be estimated in a suitable way. Appropriate estimation methods include spectral subtraction (estimating the auto power spectral density of the noise), minimum mean square error short-time spectral amplitude (MMSE STSA) estimator or MMSE log-SA estimator or a speech pause detector, for example.

[0072] It is also possible to estimate the short-time auto power spectral density of the noise signal with the help of the coherence between two or more microphones. In a second step, the estimated short-time auto power spectral density of the noise signal may be used to estimate the absolute value of the most probable Fourier coefficient (using, for example, a spectral subtraction algorithm or an MMSE log-SA estimator) and to reconstruct the absolute value of the spectrum of the speech signal. For the multi-channel noise reduction, one estimates that the coherence or the cross power spectral density of the noise signals received by the microphones is vanishing. In the case of two microphones, for example, the microphone signal has the form:

$$x_1[k] = \sum_{i=0}^{\infty} h_1(i)s[k-i] + n_1[k] \text{ and } x_2[k] = \sum_{i=0}^{\infty} h_2(i)s[k-i] + n_2[k],$$

wherein $h_1(i)$ and $h_2(i)$ are the impulse responses representing the acoustic transfer between the source of speech and the microphones. Both parts of the speech signal filtered in this way are superimposed with the uncorrelated noise signals $n_1[k]$ and $n_2[k]$.

[0073] Since $\Phi_{n_1 n_2}(\omega) = 0$ and assuming that the Fourier transforms ($H_1(\omega)$ and $H_2(\omega)$) of the impulse responses of the acoustic transfer ($h_1(i)$ and $h_2(i)$) obey $|H_1(\omega)| = |H_2(\omega)|$, one obtains for the short-time auto power spectral density

$$\hat{\Phi}_{nn}(\kappa, \nu) = P \left\{ \sqrt{\overline{E}\{|X_1(\kappa, \nu)|^2\}} \overline{E}\{|X_2(\kappa, \nu)|^2\} - \overline{E}\{|X_1(\kappa, \nu)X_2(\kappa, \nu)|^2\} \right\}$$

wherein

$$P\{x\} = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}.$$

[0074] The corresponding Wiener filter, then, has the form

$$W(\omega) = \frac{|\Phi_{x_1 x_2}(\omega)|}{\Phi_{x_1 x_1}(\omega)} = \frac{\Phi_{xx}(\omega)|H_1(\omega)|^2}{\Phi_{x_1 x_1}(\omega)}.$$

[0075] In Fig. 1, the adaption of the post-filter $w(k, i)$ - k being the time index and i denoting the coefficient within the impulse response - is performed in the time domain, for example, with the help of the LMS algorithm. The background Wiener filters $w_1(k, i), \dots, w_4(k, i)$ are minimize the error signals $e_1[k], \dots, e_4[k]$ such that, for example, the filter $w_4(k, i)$ tends

towards the frequency response

$$W_4(\omega) = \frac{\Phi_{d, \tilde{x}_4}(\omega)}{\Phi_{\tilde{x}_4, \tilde{x}_4}(\omega)},$$

wherein

$$\begin{aligned} \Phi_{d, \tilde{x}_4}(\omega) &= E \left\{ \frac{\tilde{X}_1(\omega) + \tilde{X}_2(\omega) + \tilde{X}_3(\omega)}{3} \tilde{X}_4^*(\omega) \right\} \\ &= E \left\{ \frac{\tilde{H}_1(\omega) + \tilde{H}_2(\omega) + \tilde{H}_3(\omega)}{3} S(\omega) (\tilde{H}_4(\omega) S(\omega))^* \right\} \\ &= \frac{\tilde{H}_1(\omega) + \tilde{H}_2(\omega) + \tilde{H}_3(\omega)}{3} \tilde{H}_4^*(\omega) E \{ |S(\omega)|^2 \}. \end{aligned}$$

[0076] The form of the other three Wiener filters is obtained by a cyclic permutation of the indices.

[0077] It is to be understood that the system is not restricted to a particular number of Wiener filters 110. Furthermore, not every Wiener filter 110 is always to be used to determine the adaptive post-filter 108. For example, one may use only the Wiener filter which uses the microphone signal of the microphone proximal to the source of speech.

[0078] Preferably, however, the adaptive post-filter 108 is determined as

$$w[k, i] = \frac{1}{4} (w_1[k, i] + w_2[k, i] + w_3[k, i] + w_4[k, i]).$$

[0079] The filter is linear-phase if the filter coefficients satisfy

$$w[k, i] = w[k, L - i].$$

[0080] Using this symmetry condition, the filter coefficients of a linear-phase post-filter (with length L) can be obtained. Accordingly, the linear-phase post-filter has twice the length of one of the background filters 110 (with length L/2). Such a linear-phase filter only modifies the amplitude spectrum of the input signal of the filter without a frequency dependent distortion of the phase spectrum.

[0081] The performance of the filter can be further improved by smoothing its frequency response. This can be achieved by weighting the filter coefficients with a window function.

[0082] The inverse filters 106 serve to compensate for the acoustic transfer function of the path between the source of speech and the microphones.

[0083] In Figure 2, the structure of a superdirective beamformer is shown. The beamformer shown in this figure performs the filtering in the frequency domain, in contrast to the case of Figure 1. If a beamformer in the frequency domain were used in Figure 1, an inverse Fourier transform is to be performed on the signals before passing the signals to the Wiener filters 110 or the pre-emphasis filter 111.

[0084] In Figure 2, the microphone array consists of M microphones 102, each yielding a signal $x_i(t)$. The signals $x_i(t)$ are transferred to the frequency domain by fast Fourier transform (FFT) means 201, resulting in a signal $X_i(\omega)$. In general, the beamforming consists of a beamsteering and a filtering. The beamsteering is responsible for the propagation time compensation. The beamsteering is performed by a steering vector

$$d(\omega) = [a_0 e^{-j2\pi f r_0}, a_1 e^{-j2\pi f r_1}, \dots, a_{M-1} e^{-j2\pi f r_{M-1}}],$$

with

$$a_n = \frac{\|q - p_{ref}\|}{\|q - p_n\|}$$

and

$$\tau_n = \frac{\|q - p_{ref}\| - \|q - p_n\|}{c},$$

wherein p_{ref} denotes the position of a reference microphone, p_n the position of microphone n , q the position of the source of sound (for example, the speaker), f the frequency and c the velocity of sound. In the far field, one has

$$a_0 = a_1 = \dots = a_{m-1} = 1.$$

[0085] According to a rule of thumb, one has the far field situation if the source of the useful signal is more than twice as far from the microphone array as the maximum dimension of the array. In Figure 2, a far field beamformer is shown since only a phase factor $e^{j\omega\tau_k}$ denoted by reference sign 202 is applied to the signals $X_k(\omega)$.

[0086] After the beamsteering, the signals are filtered by superdirective filters 203 that are filters in the frequency domain. The filtered signals are summed yielding a signal $Y(\omega)$. After an inverse fast Fourier transform (IFFT) by means 204, the resulting signal $y[k]$ is obtained.

[0087] The optimal filter coefficients $A_i(\omega)$ may be computed according to

$$A_i(\omega) = \frac{\Gamma(\omega)^{-1} d(\omega)}{d(\omega)^H \Gamma(\omega)^{-1} d(\omega)},$$

wherein the superscript H denotes Hermitian transposing and $\Gamma(\omega)$ is the complex coherence matrix

$$\Gamma(\omega) = \begin{pmatrix} 1 & \Gamma_{x_1 x_2}(\omega) & \dots & \Gamma_{x_1 x_M}(\omega) \\ \Gamma_{x_2 x_1} & 1 & \dots & \Gamma_{x_2 x_M}(\omega) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma_{x_M x_1}(\omega) & \Gamma_{x_M x_2}(\omega) & \dots & 1 \end{pmatrix},$$

the entries of which are the coherence functions that are defined as the normalized cross-power spectral density of two signals

$$\Gamma_{x_i x_j}(\omega) = \frac{P_{x_i x_j}(\omega)}{\sqrt{P_{x_i x_i}(\omega) P_{x_j x_j}(\omega)}}.$$

[0088] Preferably, the beamsteering is separated from the filtering step which reduces the steering vector in the design equation for the filter coefficients $A_i(\omega)$ to the unity vector

$$d(\omega) = (1, 1, \dots, 1)^T.$$

(The superscript T denotes transposing.)

[0089] In the case of an isotropic noise field in three dimensions (diffuse noise field), the coherence is given by

$$\Gamma_{x_i x_j}(\omega) = si\left(\frac{2\pi f d_{ij}}{c}\right) e^{-j \frac{2\pi f d_{ij} \cos \Theta_0}{c}},$$

with

$$si(x) = \frac{\sin x}{x}$$

and wherein d_{ij} denotes the distance between microphones i and j and Θ_0 is the angle of the main receiving direction of the microphone array or the beamformer.

[0090] The above described design rule for computing the optimal filter coefficients $A_i(\omega)$ for a homogenous diffuse noise field is based on the assumption that the microphones are perfectly matched, i.e. point-like microphones having exactly the same transfer function. In practice, therefore, a so-called regularized filter design may be used to adjust the filter coefficients. To achieve this, a scalar (the regularization parameter μ) is added at the main diagonal of the cross-correlation matrix. In a slightly modified version, all elements of the coherence matrix not on the main diagonal are divided by $(1 + \mu)$:

$$\overline{\Gamma_{x_i x_j}(\omega)} = \frac{\Gamma_{x_i x_j}(\omega)}{1 + \mu} = \frac{si\left(\frac{2\pi f d_{ij}}{c}\right)}{1 + \mu} e^{-j \frac{2\pi f d_{ij} \cos \Theta_0}{c}}, \quad \forall i \neq j.$$

[0091] Alternatively, the regularization parameter μ may be introduced into the equation for computing the filter coefficients:

$$A_i(\omega) = \frac{(\Gamma(\omega) + \mu I)^{-1} d}{d^T (\Gamma(\omega) + \mu I)^{-1} d}$$

wherein I is the unity matrix. For convenience, in the following, the second approach where the regularization parameter

is part of the filter equation will be discussed in more detail. It is to be understood, however, that the first approach is equally suitable.

[0092] Before discussing the superdirective beamformer in more detail, some characteristic quantities of a microphone array are to be defined. The directional diagram or response pattern ($\Psi(\omega, \Theta)$) of a microphone array characterizes the sensitivity of the array as a function of the direction of incidence Θ for different frequencies.

[0093] A measure to describe the directivity of an array is the so-called gain that does not depend on the angle of incidence Θ . The gain is defined as the sensitivity of the array in the main direction of incidence with respect to the sensitivity for omnidirectional incidence.

[0094] The Front-To-Back-Ratio (FBR) indicates the sensitivity in front receiving direction compared to the back.

[0095] The white noise gain (WNG) describes the ability of the array to suppress uncorrelated noise, for example, the inherent noise of the microphones. The inverse of the white noise gain is the susceptibility $K(\omega)$,

[0096] The susceptibility $K(\omega)$ describes the array's sensitivity to defective parameters. It is often preferred that the susceptibility $K(\omega)$ of the array filters $A_i(\omega)$ does not exceed an upper bound $K_{\max}(\omega)$. The selection of this upper bound may be dependent on the relative error $\Delta^2(\omega, \Theta)$ of the microphones and, for example, on requirements regarding the directional diagram $\Psi(\omega, \Theta)$. The relative error $\Delta^2(\omega, \Theta)$, in general, is the sum of the mean square error of the transfer properties of all microphones $\varepsilon^2(\omega, \Theta)$ and the Gaussian error with zero mean of the microphone positions $\delta^2(\omega)$.

[0097] Defective array parameters may also disturb the ideal directional diagram; the corresponding error can be given by $\Delta^2(\omega, \Theta)K(\omega)$. If one requires that the deviations in the directional diagram do not exceed an upper bound of $\Delta\Psi_{\max}(\omega, \Theta)$, one obtains for the maximum susceptibility:

$$K_{\max}(\omega, \Theta) = \frac{\Delta\Psi_{\max}(\omega, \Theta)}{\varepsilon^2(\omega, \Theta) + \delta^2(\omega)}$$

[0098] It is to be noted that in many cases the dependence on the angle Θ can be neglected.

[0099] In practice, the error in the microphone transfer functions $\varepsilon(\omega)$ has a higher influence on the maximum susceptibility $K_{\max}(\omega)$ and, thus, also on the maximum possible gain $G(\omega)$ than the error $\delta^2(\omega)$ in the microphone positions. In other words, the defective transfer functions are mainly responsible for the limitation of the maximum susceptibility.

[0100] A higher mechanical precision to reduce the position deviations of the microphones is only sensible up to a certain point since the microphones usually are modeled as being point-like, which is not true in reality. Thus, one can fix the positioning errors $\delta^2(\omega)$ to a specific value, even if a higher mechanical precision could be achieved. For example, one can take $\delta^2(\omega)=1\%$ which is quite realistic. The error $\varepsilon(\omega)$ can be derived from the frequency depending deviations of the microphone transfer functions.

[0101] To compensate the above-mentioned errors, inverse filters may be used to adjust the individual microphone transfer functions to a reference transfer function. Such a reference transfer function can be the transfer function of one microphone out of the array or, for example, the mean of all measured transfer functions. In case of the first possibility, only $M-1$ inverse filters (M being the number of microphones) are to be computed and implemented.

[0102] In general, the transfer functions are not minimal phase, thus, a direct inversion would yield instable filters. Usually, one inverts only the minimum phase part of the transfer function (resulting in a phase error) or one inverts the ideal (non-minimum phase) filter only approximately. In the following, the approximate inversion with the help of an FXLMS (filtered X least mean square) or the FXNLMS (filtered X normalized least mean square) algorithm will be described.

[0103] After computing of the inverse filters, they may be coupled with the superdirective filters $A_i(\omega)$ such that, in the end, only one filter per viewing direction and microphone is to be implemented.

[0104] The FXLMS or the FXNLMS algorithm is described with reference to Figure 3. The error signal $e[n]$ at time n is calculated according to

$$\begin{aligned} e[n] &= d[n] - y[n] \\ &= (\mathbf{p}^T[n] \mathbf{x}[n]) - (\mathbf{w}^T[n] \mathbf{x}'[n]) \\ &= (\mathbf{p}^T[n] \mathbf{x}[n]) - (\mathbf{w}^T[n] (\mathbf{s}^T[n] \mathbf{x}[n])) \end{aligned}$$

with the input signal vector

$$\mathbf{x}[n] = [x[n], x[n-1], \dots, x[n-L+1]]^T$$

wherein L denotes the filter length of the inverse filter $W(z)$. The filter coefficient vector of the inverse filter has the form

$$\mathbf{w}[n] = [w_0[n], w_1[n], \dots, w_{L-1}[n]]^T,$$

the filter coefficient vector of the reference transfer function $P(z)$

$$\mathbf{p}[n] = [p_0[n], p_1[n], \dots, p_{L-1}[n]]^T$$

and the filter coefficient vector of the n -th microphone transfer function $s(z)$

$$\mathbf{s}[n] = [s_0[n], s_1[n], \dots, s_{L-1}[n]]^T.$$

[0105] The update of the filter coefficients of $w[n]$ is performed iteratively, i.e. at each time step n , whereby the filter coefficient $w[n]$ are computed such that the instantaneous squared error $e^2[n]$ is minimized. This can be achieved, for example, by using the LMS algorithm:

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \mu \mathbf{x}'[n] e[n]$$

or by using the NLMS algorithm

$$\mathbf{w}[n+1] = \mathbf{w}[n] + \frac{\mu}{\mathbf{x}'[n]^T \mathbf{x}'[n]} \mathbf{x}'[n] e[n]$$

wherein μ characterizes the adaption steps and

$$\mathbf{x}'[n] = [x'[n], x'[n-1], \dots, x'[n-L+1]]^T$$

denotes the input signal vector filtered by $S(z)$.

[0106] In general, the susceptibility increases with decreasing frequency. Thus, it is preferred to adjust the microphone transfer functions depending on frequency, in particular, with a high precision for low frequencies. To achieve a high precision of the inverse filters, the FIR filters, for example, are to be very long in order to obtain a sufficient frequency resolution in the desired frequency range. This means that the expenditure, in particular, regarding the memory, increases rapidly. When using a reduced sampling frequency of, for example, $f_a = 8\text{kHz}$, the computing time does not impose a severe limitation. A suitable frequency depending adaption of the transfer functions can be achieved by using short WFIR filters (warped filters).

[0107] One possible iterative method to design the filters $A_f(\omega)$ with predetermined susceptibility goes as follows:

1. Set $\mu(\omega) = 1$.
2. Determine the transfer functions of the filters $A_i(\omega)$ and the resulting susceptibilities $K(\omega)$ according to the equations:

$$A_i(\omega) = \frac{(\Gamma(\omega) + \mu I)^{-1} d}{d^T (\Gamma(\omega) + \mu I)^{-1} d}$$

and

$$K(\omega) = \frac{1}{WNG(\omega)} = \frac{A(\omega)^H A(\omega)}{|A(\omega)^H d(\omega)|}.$$

3. If the susceptibility $K(\omega)$ is larger than the maximum susceptibility ($K(\omega) > (K_{\max}(\omega))$), increase μ in the following step, otherwise, decrease μ .
4. Repeat steps 2 and 3 until the susceptibility $K(\omega)$ is sufficiently close to the predetermined value $K_{\max}(\omega)$. The iteration is to break off if μ becomes smaller than a lower limit of, for example, $\mu_{\min} = 10^{-8}$. Such a termination criterion is mainly necessary for high frequencies $f \geq c/(2d_{\text{mic}})$.

[0108] Of course, there are other possibilities to compute the filters $A_i(\omega)$. For example, one can use a fixed parameter μ for all frequencies. This simplifies the computation of the filter coefficients. It is to be noted that the above iterative method is not used for a real time adaption of the filter coefficients during operation.

[0109] A realization of the beamforming filters in the time domain - as in the system of Figure 1 - is described with reference to Figure 4. Signals are recorded by microphones 102. A near field beamsteering 104 is performed using gain factors v_k 401 to compensate for the amplitude differences and time delays τ_k 402 to compensate for the propagation time differences of the microphone signals $x_k[l]$. The realization of the superdirective beamforming is achieved using the filters (preferably, FIR filters) $\alpha_k(i)$ indicated by reference sign 403.

[0110] The impulse responses $\alpha_1(i), \dots, \alpha_M(i)$ can be determined as follows:

1. Determine the frequency responses $A_i(\omega)$ according to the above equation.
2. To obtain real valued impulse responses $\alpha_1(i), \dots, \alpha_M(i)$, chose the frequency responses above half of the sampling

frequency to $(A_i(\omega) = A_i^*(\omega_A - \omega))$ with ω_A denoting the sampling angular frequency.

3. Transfer these frequency responses to the time domain using an IFFT yielding the desired FIR filter coefficients $\alpha_1(i), \dots, \alpha_M(i)$.
4. Applying a window function, for example, a Hamming window, to the FIR filter coefficients $a_1(i), \dots, \alpha_M(i)$.

[0111] As can be seen in Figure 4, in contrast to the beamforming in the frequency domain as described above, the microphone signals are directly processed using the beamsteering 104 in the time domain. The beamsteering 104 is followed by the FIR filtering 403. After summing the filtered signals, a resulting enhanced signal $y[k]$ is obtained.

[0112] Depending on the distance between speaker and microphone array, on the distance between the microphones themselves, and on the sampling frequency f_a , more or less propagation or transit time between the microphone signals is to be compensated. The following equation is to be taken into account:

$$\Delta_{\max} = \frac{d_{\text{mic}} f_a}{c}.$$

[0113] The higher the sampling frequency f_a or the higher the distance between adjacent microphones, the more transit time $\Delta_{\max}$ (in taps of delay) is to be compensated for. The number of taps increases also if the distance between speaker and microphone arrays is decreased. In the near field, more transit time is to be compensated for than in the far field. It turns out that an array in endfire orientation is less sensitive to a defective transit time compensation $\Delta_{\max}$

than an array in broad-side orientation.

[0114] In a vehicle, the average distance between the speaker, in particular, its head, and the array is about 50cm. Due to a movement of the head, this distance can change by about +/- 20cm. If a transit time error of 1 tap is acceptable, the distance between the microphones in broad-side orientation with a sampling frequency of $f_a = 8\text{kHz}$ should be smaller than about d_{mic_max} (broad -side) $\cong 5\text{cm}$. With the same conditions, the maximum distance between the microphones in

endfire orientation may be about $d_{mic_max}(\text{endfire}) \cong 20\text{cm}$.
[0115] On the other hand, having a distance between the microphones of about 5cm, it turns out that a sampling frequency of $f_a = 16\text{kHz}$ provides excellent results for an endfire orientation whereas in broad-side orientation, only a sampling frequency of $f_a = 8\text{kHz}$ can be used without adaptive beamsteering. In other words, in endfire orientation, the sampling frequency or the distance between the microphones can be chosen much higher than in the broad-side case, thus, resulting in an improved beamforming.

[0116] In this context, it is to be pointed out that the larger the distance between the microphones, the sharper the beam, in particular, for low frequencies. A sharper beam at low frequencies increases the gain in this range which is important for vehicles where the noise is mostly a low frequency noise.

[0117] However, the larger the microphone distance, the smaller the usable frequency range according to the spatial sampling theorem

$$f \leq \frac{c}{2d_{mic}}.$$

[0118] A violation of this sampling theorem has the consequence that at higher frequencies, large grating lobes appear. These grating lobes, however, are very narrow and deteriorate the gain only slightly. The maximum microphone distance that can be chosen depends not only on the lower limiting frequency for the optimization of the directional characteristic, but also on the number of microphones and on the distance of the microphone array to the speaker. In general, the larger the number of microphones, the smaller their maximum distance in order to optimize the Signal-To-Noise-Ratio (SNR). For a distance between array and speaker of 50cm, the microphone distance, preferably, is about $d_{mic} = 40\text{cm}$ with two microphones ($M = 2$) and about $d_{mic} = 20\text{cm}$ for $M = 4$.

[0119] A further improvement of the directivity, and, thus, of the gain, can be achieved by using unidirectional microphones instead of omnidirectional ones; this will be discussed in more detail below.

[0120] Figures 5A and 5B show preferred arrangements of microphone arrays in a vehicle. In general, the distance between the microphone array and the speaker should be as small as possible.

[0121] According to a first embodiment (Figure 5A), each speaker 501 may have its own microphone array comprising at least two microphones 102. The microphone arrays may be provided at different locations, for example, within the headliner, dashboard, pillar, headrest, steering wheel, compartment door, visor or (driving) mirror. An arrangement within the roof is also a preferred possibility that is, however, not suitable for the case of a cabriolet. Both microphone arrays for each speaker are in endfire orientation.

[0122] In an alternative embodiment (Figure 5B), one microphone array is used for two neighboring speakers. In both embodiments, preferably, directional microphones, in particular, having a cardioid characteristic, may be used.

[0123] In the embodiment of Figure 5B, the microphone array may be mounted within the mirror. Such a linear microphone array may be used for both the driver and the front seat passenger. A costly mounting of the microphones in the roof can be avoided. Furthermore, the array may be mounted in one piece, which ensures a high mechanical precision. Due to the adjustment of the mirror, the array would always be correctly oriented.

[0124] Figure 6A shows a top view on a (driving) mirror 601 of a car with three microphones in two alternative arrangements. According to the first alternative, two microphones 602 and 603 are located in the center of the mirror in endfire orientation with respect to the driver and, preferably, have a distance d_{mic} of about 5cm between each other. The microphones 603 and 604 are in endfire orientation with respect to the front seat passenger and have a distance of about 10cm between each other. Since the microphone 603 is used for both arrays, a cheap handsfree system can be provided.

[0125] All three microphones may be directional microphones, preferably having a cardioid characteristic, for example, a hypercardioid characteristic. Alternatively, microphones 602 and 604 are directional microphones, whereas microphone 603 is an omnidirectional microphone which further reduces the costs. If all three microphones are directional microphones, preferably, microphones 602 and 603 are directed towards the driver.

[0126] Due to the larger distance between microphones 603 and 604 than between microphones 602 and 603, the front seat passenger beamformer has a better SNR at low frequencies.

[0127] According to an alternative embodiment, the microphone array for the driver consists of microphones 602' and 603' located at the left side of the mirror. In this case, the distance between this microphone array and the driver would

be increased, thus, decreasing the performance. On the other hand, the distance between microphone 603' and 604 would be about 20cm, which yields a better gain for the front seat passenger at low frequencies.

[0128] A variant of two microphone arrays with improved precision is shown in Figure 6B. Also in this case, all microphones may be directional microphones, microphones 602 and 603 being directed to the driver, microphones 604 and 605 being directed to a front seat passenger. In this example, the microphone array for the front seat passenger comprises the three microphones 603, 604 and 605, which increases the gain considerably.

[0129] It is to be noted that these arrangements are only examples that may be varied by changing the position and number of the microphones. In particular, an arrangement may be optimized with regard to a specific vehicular cabin.

[0130] Figure 7 illustrates a microphone array comprising three subarrays 701, 702, and 703, each subarray consisting of five microphones. Within each subarray 701, 702, and 703, the microphones are equidistantly arranged. In the total array 704, the distances are no longer equal. As can be seen in this figure, some microphones are used for different arrays, therefore, for the total array, only 9 microphones and not $3.5 \cdot 15$ microphones are necessary.

[0131] In this figure, it is further indicated that the different subarrays are used for different frequency ranges. The resulting directional diagram is then built up of the directional diagrams of each subarray for the respective frequency range. For the special case of Figure 7, subarray 701 with $d_{mic} = 5cm$ is used for the frequency band of 1400 - 3400 Hz, subarray 702 with $d_{mic} = 10cm$ for the frequency band of 700 - 1400 Hz, and subarray 703 with $d_{mic} = 20cm$ for the band of frequencies smaller than 700 Hz. A lower limit of this frequency band may be imposed, for example, by the lowest frequency of the telephone band (the frequencies used in telephone applications) which, presently, is 300 Hz in most cases.

[0132] An improved directional characteristic may be obtained if the superdirective beamformer is designed as general sidelobe canceller (GSC). In this structure, at least one filter can be saved. Such a superdirective beamformer in GSC structure is shown in Figure 8. The GSC structure is to be implemented in the frequency domain, thus, an FFT 201 is applied to the incoming signals $x_k(t)$ from microphones 102. Before the general sidelobe cancelling, a time alignment using phase factors $e^{j\omega\tau_k}$ has to be performed (in this figure, a far field beamsteering is shown). If a GSC beamformer is used in the handsfree system of Figure 1, for example, again, an inverse Fourier transform is to be performed before passing the signal to the Wiener filters 110 or the pre-emphasis filter 111.

[0133] In Figure 8, $\mathbf{X}$ denotes a vector comprising all time aligned input signals $X_i(\omega)$. $\mathbf{A}^c$ is a vector comprising all frequency independent filter transfer functions A_i that are necessary to observe the constraints in viewing direction; $\mathbf{H}$ is the vector of the transfer functions performing the actual superdirectivity; and $\mathbf{B}$ is the so-called blocking matrix projecting the input signals in $\mathbf{X}$ onto the "noise plane". The signal $Y_{DS}(\omega)$ denotes the output signal of the delay and sum beamformer, $Y_{BM}(\omega)$ the resulting output signal of the blocking branch, $Y_{SD}(\omega)$ the output signal of the superdirective beamformer $x_i(t)$, and $X_i(\omega)$ the input signals in the time and frequency domain that are not yet time aligned, and $Y_i(\omega)$ the output signals of the blocking matrix that ideally should block completely the desired or useful signal within the input signals. The signals $Y_i(\omega)$ ideally only comprise the noise signals.

[0134] In addition to the superdirective output signal, a GSC structure also yields a delay and sum beamformer signal and a blocking output signal. The number of filters that can be saved using the GSC depends on the choice of the blocking matrix. Usually, a Walsh-Hadamard blocking matrix is preferred instead of a Griffiths-Jim blocking matrix since more filters can be saved with a Walsh-Hadamard blocking matrix. Unfortunately, the Walsh-Hadamard blocking matrix can only be given for arrays consisting of $M = 2^n$ microphones.

[0135] In principle, a blocking matrix should have the following properties:

1. It is a $(M-1) \cdot M$ -Matrix.
2. The sum of the values within one row vanishes.
3. The matrix is of rank $M-1$.

[0136] A Walsh-Hadamard blocking matrix for $n = 2$ has the following form

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}.$$

[0137] According to an alternative embodiment, a blocking matrix according to Griffiths-Jim can be used which has the general form

$$\mathbf{B} = \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & -1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & -1 \end{bmatrix}.$$

[0138] The upper branch of the GSC structure is a delay and sum beamformer with the transfer functions

$$\mathbf{A}^c = \left[\underbrace{\frac{1}{M}, \frac{1}{M}, \dots, \frac{1}{M}}_M \right]^T.$$

[0139] The computation of the filter coefficients of a superdirective beamformer in GSC structure is slightly different compared to the conventional superdirective beamformer. The transfer functions $H_i(\omega)$ are to be computed as

$$H_i(\omega) = (\mathbf{B}\Phi_{nn}(\omega)\mathbf{B}^H)^{-1}(\mathbf{B}\Phi_{nn}(\omega)\mathbf{A}^c),$$

wherein $\mathbf{B}$ is the blocking matrix and $\Phi_{nn}(\omega)$ the matrix of the cross-correlation power spectrum of the noise. In the case of a homogenous noise field, $\Phi_{nn}(\omega)$ may be replaced by the time aligned coherence matrix of the diffuse noise field, as discussed above.

[0140] A regularization and the iterative design with predetermined susceptibility may be performed in the same way as above.

[0141] All previously discussed filter designs only assume that the noise field is homogenous and diffuse. These designs may be generalized by excluding a region around the main receiving direction Θ_0 when determining the homogenous noise field. In this way, mainly the Front-To-Back-Ratio may be optimized. This is illustrated in Figure 9 with microphones 102 where a sector of $\pm\delta$ is excluded. The computing of the two-dimensional diffuse (cylindrically isotropic), homogenous noise field can be performed using the new design parameter δ :

$$\Gamma(\omega, \Theta_0, \delta) = \frac{1}{2(\pi - \delta)} \int_{\Theta_0 - \delta}^{\Theta_0 + \delta} e^{j\left(\frac{2\pi d_y \cos \Theta}{c}\right)} d\Theta e^{-j\left(\frac{2\pi d_y \cos \Theta_0}{c}\right)}, \quad i, j \in [1, \dots, M]$$

[0142] This method may also be generalized to the three-dimensional case. Then, in addition to the parameter δ being responsible for the azimuth, a further parameter ρ is to be introduced for the elevation angle. This yields an analog equation for the coherence of the homogeneous diffuse 3D noise field.

[0143] A superdirective beamformer based on an isotropic noise field is particularly useful for a handsfree system which is to be installed later in a vehicle. This is the case, for example, if the handsfree system is installed in the vehicle by the user itself. On the other hand, an MVDR beamformer may be relevant if there are specific noise sources at fixed relative positions or directions with respect to the position of the microphone array. In this case, the handsfree system can be adapted to a particular vehicular cabin by adjusting the beamformer such that its zeros point into the direction of specific noise sources. For example, such a noise source may be formed by a loudspeaker or a fan. Preferably, a handsfree system with MVDR beamformer is already installed during manufacture of the vehicle.

[0144] The typical distribution of noise or noise sources in a particular vehicular cabin can be determined by performing corresponding noise measurements under appropriate conditions (e.g., driving noise with and/or without loudspeaker and/or fan noise). The measured data are used for the design of the beamformer. It is to be noted that also in this case, no further adaption is performed during operation of the handsfree system.

[0145] Alternatively, if the relative position of a noise source is known, the corresponding superdirective filter coefficients can also be determined theoretically.

[0146] As already stated above, the use of directional microphones further improves the signal enhancement. Figure 10 shows a superdirective beamformer with directional microphones 1001. In this figure, each directional microphone 1001 is depicted by its equivalent circuit diagram. In these circuit diagrams, d_{DMA} denotes the (virtual) distance of the two omnidirectional microphones composing the first order pressure gradient microphone in the circuit diagram. T is the (acoustic) delay line fixing the characteristic of the directional microphone and EQ_{TP} is the equalizing low path filter yielding a frequency independent transfer behavior in viewing direction.

[0147] In practice, these circuits and filters may be realized purely mechanically by taking an appropriate mechanical directional microphone. Again, the distance between the directional microphones is d_{mic} . In Figure 10, the whole beamforming is performed in the time domain. A near field beamsteering 104 is applied to the signals $x_n[l]$ coming from the microphones and being filtered by the equalizing filter EQ_{TP} . The gain factors v_n compensate for the amplitude differences and the delays τ_n for the transit time differences of the signals. The FIR filters $\alpha_n[l]$ realize the superdirectivity in the time domain.

[0148] Mechanical pressure gradient microphones have a high quality and yield, in particular, using a hypercardioid characteristic, an excellent array gain. The use of directional microphones results in an excellent Front-to-Back-Ratio as well.

[0149] An example for another preferred embodiment of a handsfree system is shown in Fig. 11. The system shown in this figure differs from the system of Fig. 1 in that an adaptive noise canceller (ANC) system 1101 is provided between the microphone array 101 and the high-pass filters 103. The ANC system is particularly useful to reduce motor harmonics in the signal.

[0150] The structure of an adaptive noise canceller is shown in Fig. 12. In this figure, a wanted signal source 1201, particularly corresponding to a speaker, and a noise source 1202 are shown. The signal entering a microphone 102 which is part of the microphone array is the sum of a wanted signal $s[k]$ and a noise signal $n_0[k]$. In addition, a noise sensor 1203 is present which is to provide a pure noise signal $n_1[k]$. In Fig. 12, the reference sensor 1203 is a microphone; in this case, such a microphone should be arranged at a place where no or almost no wanted signal is to be recorded. The output signal $y[k]$ of the adaptive filter 1204 is subtracted from the output signal of the microphone 102. The input signal $n_1[k]$ for the adaptive filter 1204 and the output signal or error signal $e[k]$ serve for adaption of the noise canceller. The noise reduction of the adaptive noise canceller depends only on the coherence of the signals of the microphone 102 and the reference sensor 1204; this coherence function in turn is depending on the distance between microphone and reference sensor.

[0151] According to a preferred alternative, the reference sensor is not an acoustic sensor. One possibility is to couple an electrical sensor with the tachometer or speed counter which is usually present in a vehicle. After determining the interrelationship between the tachometer signal and the motor noise, the latter may be subtracted from the microphone signal via the adaptive noise canceller. Such an embodiment is shown in Fig. 11 where a tachometer 1102 is coupled to the ANC 1101.

[0152] The ANC need not be placed directly behind the microphone array 101. According to a preferred alternative, the ANC may be used to filter the output signal $x[k]$ of the superdirective beamformer. In this case, the ANC is to be placed between the summing circuit 107 and the adaptive post-filter 108.

[0153] A further noise reduction with the help of an ANC system can be achieved by using additional - acoustical or non-acoustical - noise sensors. A corresponding embodiment is shown in Fig. 13. In this embodiment, the ANC system 1304 is used particularly to suppress signals coming from a loudspeaker 1301, for example, emitting a far end signal 1302. The ANC system is able to create a so-called area of silence around the noise sensor or noise sensors. If the microphone array 101 is located in the vicinity of the near speaker, the whole array 101 or one of its microphones 102 may be used as noise sensor. Alternatively, one or more acoustical (1203) and/or non-acoustical (1102) noise sensors are to be installed. Ideally, an acoustical ANC can provide a noise reduction for both the near and far end speaker.

[0154] As can be seen in Fig. 13, an additional acoustic echo canceller (AEC) system 1303 may also be provided. Such an AEC system is optional, but allows a suppression of reverberation. Preferably, each microphone 102 is provided with an individual AEC filter. Alternatively, an AEC filter may be placed between the summing circuit and the adaptive post-filter. In such an alternative embodiment, an ANC may be placed between the summing circuit and the AEC.

[0155] Advantageously, the AEC system used for this invention comprises a conventional AEC filter and an integrated frequency selected echo attenuation which acts as a residual echo suppression (RES) algorithm. A preferred embodiment of such a system is shown in Fig. 14. It comprises a conventional AEC filter 1303 that filters the far end signal 1302. The adaption of the conventional AEC filter 1303 is performed using the signal 1401, e.g. the output signal of an electrical ANC or of one of the microphones. Furthermore, an echo shaping means 1402 is provided. This echo shaping means has the form of an adaptive FIR filter with coefficient vector $H[k]$ that filters the compensated signal $e[k]$. The coefficient vector $H[k]$ of the adaptive filter is taken in each sampling step from another adaptive FIR filter with coefficient vector $H1[k]$. Preferably, the filter with coefficient vector $H1[k]$ is a linear-phase filter of low order. The echo shaping means

further comprises a delay element T_{H1} . The adaption algorithm is based on the power difference between $y[k]$ and $e[k]$ by filtering according to

$$z[k] = a[k]y[k] + (1 - a[k])e[k],$$

wherein the compensated signal $e[k]$ is the reference signal. The resulting signal $z[k]$ depends on the time varying factor $\alpha[k]$. In case the far end speaker is active, the output signal $y[k]$ of the ANC system is dominant ($\alpha[k]$ close to 1) at the input of the adaptive filter with coefficient vector $H1[k]$. Then, the adaptive filter with coefficient vector $H1[k]$ can reduce the error signal $\tilde{E}[k]$ only by suppressing the signal of the far end speaker. In this case, the near speaker and local noise signals will not be attenuated by the echo shaping means. It is to be noted that the echo shaping algorithm is frequency selective.

[0156] The system shown in Fig. 15 is able to process speech signals from two different positions (for example, from the driver and the front seat passenger in a car). The microphone array 101 has a directional diagram with two preferred directions. For example, directional microphones may be used and/or the microphones may be arranged in a suitable way. One or several ANC or AEC filters can provide an estimation of the noise level present in the microphone signals that may be used in the dynamic volume control (DVC) 1501 to vary the volume of the far end speech signal 1511 in dependence of the noise level. The system comprises a beamformer for both wanted signal sources each comprising a beamsteering means 1502 and 1504 and beamformer filters 1503 and 1505. Following each beamformer, adaptive post-filters 1506 and 1507 are arranged which, in turn, are directly connected to non-adaptive post-filters 1508 and 1509.

[0157] The output signals of the non-adaptive post-filters are fed to a unit 1510 comprising two voice activity detectors and a switch control that generates a weighting factor A for combining both signals. According to a possible example

for such a unit 1510, each of the signals $s\hat{s}_1[k]$ and $s\hat{s}_2[k]$ can be processed by a low-pass filter. Then, for each of the filtered signals, the contained noise signal or its level is estimated using, for example, a minimum statistics. The noise signal level is subtracted from the corresponding filtered signal level. The resulting signal levels are compared to a threshold value. Depending on this comparison of both signal levels, the weighting factor A is determined. For example, a possible weighting can be determined as follows. If both signal levels are larger than the threshold value, both signals are equally weighted. If one of the signal levels is larger than the threshold value and the other is smaller than the threshold value, the larger signal is weighted by a factor of 1 and the other is fully suppressed (weighting factor 0). If both signal levels are smaller than the threshold value, the signal stemming from the direction of the driver's seat is weighted by a factor of 1 and the other signal is fully suppressed.

[0158] The combined signal may be subject to an additional post processing 1512, for example, a residual echo suppression (RES). For a RES, the combined signal is weighted by a spectral short time gain in the frequency domain, wherein the gain factor depends on the spectrum of the far end speech signal.

Claims

1. Handsfree system for use in a vehicle comprising a microphone array (101) with at least two microphones (102), a signal processing means, and an adaptive post-filter (108),
wherein the signal processing means comprises a beamformer having an input connected to the at least two microphones and an output connected to the input of the adaptive post-filter, the beamformer comprising a beamsteering means and a summing means,
wherein the adaptive post-filter is a linear-phase Wiener filter.
2. Handsfree system according to claim 1, wherein the adaptive post-filter is a filter in the time domain.
3. Handsfree system according to one of the preceding claims, wherein the signal processing means further comprises at least two adaptive filters (110) having an input connected to the output of the beamsteering means (104) and an output connected to the adaptive post-filter, wherein the at least two adaptive filters are configured to determine adaptive filter parameters for the adaptive post-filter.
4. Handsfree system according to claim 3, wherein for each of the at least two microphones, an adaptive filter is provided having an input connected to the output of the beamsteering means.

5. Handsfree system according to claim 3 or 4 , wherein an input of each of the at least two adaptive filters is further connected to the output of the beamformer.
- 5 6. Handsfree system according to one of the preceding claims, wherein the signal processing means further comprises a pre-emphasis filter (111), in particular, comprising a pre-whitening filter, having an input connected to an output of the adaptive post-filter and/or a pre-emphasis filter, in particular, comprising a pre-whitening filter, having an input connected to the output of the beamsteering means and an output connected to the at least two adaptive filters.
- 10 7. Handsfree system according to one of the preceding claims, wherein the signal processing means further comprises an inverse filter (106), particularly a warped inverse filter.
8. Handsfree system according to one of the preceding claims, wherein the signal processing means further comprises a non-adaptive post-filter (109) having an input connected to an output of the adaptive post-filter.
- 15 9. Handsfree system according to one of the preceding claims, wherein the signal processing means further comprises an adaptive noise canceller (ANC) (1101) and/or an acoustic echo canceller (AEC) (1303).
10. Handsfree system according to one of the preceding claims, wherein the beamformer is a non-adaptive beamformer and/or a superdirective beamformer and/or a minimum variance distortionless response (MVDR) beamformer.
- 20 11. Handsfree system according to one of the preceding claims, wherein the microphone array comprises at least two microphones being arranged in endfire orientation with respect to a first position.
- 25 12. Handsfree system according to claim 11, wherein the microphone array comprises at least two microphones being arranged in endfire orientation with respect to a second position.
13. Handsfree system according to claim 12, wherein the at least two microphones in the first endfire orientation and the at least two microphones in the second endfire orientation have a microphone in common.
- 30 14. Handsfree system according to one of the preceding claims, wherein the signal processing means comprises at least two beamformers.
15. Handsfree system according to claim 14, further comprising a voice activity detector (VAD) and/or a switch control means (1510).
- 35 16. Handsfree system according to one of the preceding claims, further comprising a residual echo suppression (RES) means and/or a dynamic volume control (DVC) (1510).
- 40 17. Handsfree system according to one of the preceding claims, wherein the microphone array comprises at least two subarrays (701, 702, 703).
18. Handsfree system according to one of the preceding claims, further comprising a frame wherein each microphone of the microphone array is arranged in a predetermined, in particular fixed, position in or on the frame.
- 45 19. Handsfree system according to one of the preceding claims, wherein at least one microphone is a directional microphone (1001), in particular, having a cardioid characteristic and/or being a differential microphone.
20. Vehicle comprising a handsfree system according to one of the preceding claims.
- 50 21. Method for noise reduction in a vehicular handsfree system, comprising receiving input signals resulting from a microphone array with at least two microphones, processing the input signals by a beamformer to provide a beamformed signal, and adaptively filtering a signal resulting from the beamformed signal by an adaptive post-filter, wherein the adaptive post-filter is a linear-phase Wiener filter.
- 55 22. Method according to claim 21, wherein the adaptive filtering is performed in the time domain.
23. Method according to claim 21 or 22, further comprising

providing at least two adaptive filters, particularly Wiener filters, wherein processing the input signals by a beamformer comprises beamsteering the input signals for providing beamsteered signals corresponding to one of the at least two microphones and summing the signals, and adaptively filtering comprises receiving and processing at least one beamsteered signal by at least one of the at least two adaptive filters to determine adaptive filter parameters for the adaptive post filter.

24. Method according to claim 23, wherein adaptively filtering further comprises receiving a signal resulting from the beamformed signal by at least one adaptive filter and wherein processing the beamsteered signal comprises determining adaptive filter parameters using the at least one beamsteered signal and the signal resulting from the beamformed signal.

25. Method according to claim 23 or 24, wherein for each beamsteered signal, an adaptive filter is provided for determining adaptive filter parameters using the beamsteered signal and the signal resulting from the beamformed signal.

26. Method according to one of the claims 23 - 25, wherein receiving at least one beamsteered signal by at least one of the at least two adaptive filters comprises processing the at least one beamsteered signal by a pre-emphasis filter, in particular, comprising a pre-whitening filter.

27. Method according to one of the claims 21 - 26, further comprising processing a signal resulting from the microphone array by an inverse filter, in particular, a warped inverse filter.

28. Method according to one of the claims 21 - 27, further comprising non-adaptively filtering a signal resulting from the adaptively filtered signal and/or processing a signal resulting from the adaptively filtered signal by a pre-emphasis filter.

29. Method according to one of the claims 21 - 28, further comprising processing a signal resulting from the microphone array, particularly resulting from the beamformed signal, by an adaptive noise canceller (ANC) and/or an acoustic echo canceller (AEC) and/or a residual echo suppression (RES) means.

30. Method according to one of the claims 21 - 29, wherein the input signals are processed by a non-adaptive and/or superdirective and/or minimum variance distortionless response (MVDR) beamformer.

31. Computer program product comprising one or more computer readable media having computer-executable instructions for performing the steps of the method according to claims 21 - 30.

Patentansprüche

1. Freisprechsystem zur Verwendung in einem Fahrzeug, umfassend ein Mikrofon-Array (101) mit wenigstens zwei Mikrofonen (102), ein Signalverarbeitungsmittel und ein adaptives Nachfilter (108), wobei das Signalverarbeitungsmittel einen Strahlformer mit einem Eingang, der mit den wenigstens zwei Mikrofonen verbunden ist, und einem Ausgang, der mit dem Eingang des adaptiven Nachfilters verbunden ist, aufweist, wobei der Strahlformer ein Strahlenkittel und ein Summiermittel umfasst; wobei das adaptive Nachfilter ein Linearphasen-Wiener-Filter ist.

2. Freisprechsystem nach Anspruch 1, wobei das adaptive Nachfilter ein Filter in der Zeitdomäne ist.

3. Freisprechsystem nach einem der vorherigen Ansprüche, wobei das Signalverarbeitungsmittel des Weiteren wenigstens zwei adaptive Filter (110) mit einem Eingang, der mit dem Ausgang des Strahlenkittels (104) verbunden ist, und einem Ausgang, der mit dem adaptiven Nachfilter verbunden ist, umfasst, wobei die wenigstens zwei adaptiven Filter ausgestaltet sind zum Bestimmen von adaptiven Filterparametern für das adaptive Nachfilter.

4. Freisprechsystem nach Anspruch 3, wobei für jedes der wenigstens zwei Mikrofone ein adaptives Filter mit einem Eingang vorgesehen ist, der mit dem Ausgang des Strahlenkittels verbunden ist.

5. Freisprechsystem nach Anspruch 3 oder 4, wobei ein Eingang eines jeden der wenigstens zwei adaptiven Filter des Weiteren mit dem Ausgang des Strahlformers verbunden ist.

6. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei das Signalverarbeitungsmittel des Weiteren umfasst: ein Vorverzerrungsfilter (111), das insbesondere ein Prewhitening-Filter umfasst, mit einem Eingang, der mit einem Ausgang des adaptiven Nachfilters verbunden ist, und/oder ein Vorverzerrungsfilter, das insbesondere ein Prewhitening-Filter umfasst, mit einem Eingang, der mit dem Ausgang des Strahlenmittels verbunden ist, und einem Ausgang, der mit den wenigstens zwei adaptiven Filtern verbunden ist.
7. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei das Signalverarbeitungsmittel des Weiteren ein Inversfilter (106), insbesondere ein Verzerrungsinversfilter (warped inverse filter), aufweist.
8. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei das Signalverarbeitungsmittel des Weiteren ein nichtadaptives Nachfilter (109) mit einem Eingang, der mit einem Ausgang des adaptiven Nachfilters verbunden ist, umfasst.
9. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei das Signalverarbeitungsmittel des Weiteren einen adaptiven Rauschbeseitiger (Adaptive Noise Canceller ANC) (1101) und/oder einen Akustikechobeseitiger (Acoustic Echo Canceller AEC) (1303) umfasst.
10. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei der Strahlformer ein nichtadaptiver Strahlformer und/oder ein superdirektiver Strahlformer und/oder ein MVDR-Strahlformer (Minimum Variance Distortionless Response MVDR, verzerrungsfreie Reaktion mit minimaler Varianz) ist.
11. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei das Mikrofon-Array wenigstens zwei Mikrofone umfasst, die in einer Längsstrahlorientierung (endfire orientation) in Bezug auf eine erste Position angeordnet sind.
12. Freisprechsystem nach Anspruch 11, wobei das Mikrofon-Array wenigstens zwei Mikrofone umfasst, die in einer Längsstrahlorientierung (endfire orientation) in Bezug auf eine zweite Position angeordnet sind.
13. Freisprechsystem nach Anspruch 12, wobei die wenigstens zwei Mikrofone in der ersten Längsstrahlorientierung und die wenigstens zwei Mikrofone in der zweiten Längsstrahlorientierung ein Mikrofon gemeinsam haben.
14. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei das Signalverarbeitungsmittel wenigstens zwei Strahlformer umfasst.
15. Freisprechsystem nach Anspruch 14, des Weiteren umfassend einen Sprachaktivitätsdetektor (Voice Activity Detector VAD) und/oder ein Schaltsteuermittel (1510).
16. Freisprechsystem nach einem der vorhergehenden Ansprüche, des Weiteren umfassend ein RES-Mittel (Residual Echo Suppression RES, Restechounterdrückung) und/oder eine dynamische Lautstärkesteuerung (Dynamic Volume Control DVC) (1510).
17. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei das Mikrofon-Array wenigstens zwei Teil-Arrays (701, 702, 703) umfasst.
18. Freisprechsystem nach einem der vorhergehenden Ansprüche, des Weiteren umfassend einen Rahmen, in dem jedes Mikrofon des Mikrofon-Arrays in einer vorbestimmten, insbesondere festen Position in oder an dem Rahmen angeordnet ist.
19. Freisprechsystem nach einem der vorhergehenden Ansprüche, wobei wenigstens ein Mikrofon ein Richtmikrofon (1001), insbesondere mit einer Kardioid-Kennkurve, und/oder ein Differenzialmikrofon ist.
20. Fahrzeug, umfassend ein Freisprechsystem nach einem der vorhergehenden Ansprüche.
21. Verfahren zur Rauschverringern in einem Fahrzeugfreisprechsystem, umfassend:
Empfangen von aus einem Mikrofon-Array mit wenigstens zwei Mikrofonen resultierenden Eingangssignalen, Verarbeiten der Eingangssignale durch einen Strahlformer zur Bereitstellung eines strahlgeformten Signals und adaptives Filtern eines aus dem strahlgeformten Signal resultierenden Signals durch ein adaptives Nachfilter,

wobei das adaptive Nachfilter ein Linearphasen-Wiener-Filter ist.

22. Verfahren nach Anspruch 21, wobei das adaptive Filtern in der Zeitdomäne vorgenommen wird.

23. Verfahren nach Anspruch 21 oder 22, des Weiteren umfassend:

Bereitstellen von wenigstens zwei adaptiven Filtern, insbesondere Wiener-Filtern, wobei:

das Verarbeiten der Eingangssignale durch einen Strahlformer ein Strahlenken der Eingangssignale zur Bereitstellung von strahlgelenkten Signalen entsprechend einem der wenigstens zwei Mikrofone und ein Summieren der Signale umfasst, und
das adaptive Filtern ein Empfangen und Verarbeiten wenigstens eines strahlgelenkten Signals durch wenigstens eines der wenigstens zwei adaptiven Filter zur Bestimmung von adaptiven Filterparametern für das adaptive Nachfilter umfasst.

24. Verfahren nach Anspruch 23, wobei das adaptive Filtern des Weiteren ein Empfangen eines aus dem strahlgeformten Signal resultierenden Signals durch wenigstens ein adaptives Filter umfasst und wobei das Verarbeiten des strahlgelenkten Signals ein Bestimmen von adaptiven Filterparametern unter Verwendung des wenigstens einen strahlgelenkten Signals und des aus dem strahlgeformten Signal resultierenden Signals umfasst.

25. Verfahren nach Anspruch 23 oder 24, wobei für jedes strahlgelenkte Signal ein adaptives Filter zum Bestimmen von adaptiven Filterparametern unter Verwendung des strahlgelenkten Signals und des aus dem strahlgeformten Signal resultierenden Signals vorgesehen ist.

26. Verfahren nach einem der Ansprüche 23 bis 25, wobei das Empfangen wenigstens eines strahlgelenkten Signals durch wenigstens eines der wenigstens zwei adaptiven Filter ein Verarbeiten des wenigstens einen strahlgelenkten Signals durch ein Vorverzerrungsfilter, das insbesondere ein Prewhitening-Filter umfasst, umfasst.

27. Verfahren nach einem der Ansprüche 21 bis 26, des Weiteren umfassend ein Verarbeiten eines aus dem Mikrofon-Array resultierenden Signals durch ein Inversfilter, insbesondere ein Verzerrungsinversfilter (warped inverse filter).

28. Verfahren nach einem der Ansprüche 21 bis 27, des Weiteren umfassend ein nichtadaptives Filtern eines aus dem adaptiv gefilterten Signal resultierenden Signals und/oder Verarbeiten eines aus dem adaptiv gefilterten Signal resultierenden Signals durch ein Vorverzerrungsfilter.

29. Verfahren nach einem der Ansprüche 21 bis 28, des Weiteren umfassend ein Verarbeiten eines aus dem Mikrofon-Array, insbesondere aus dem strahlgeformten Signal resultierenden Signals durch einen adaptiven Rauschbeseitiger (Adaptive Noise Canceller ANC) und/oder einen Akustikechobeseitiger (Acoustic Echo Canceller AEC) und/oder ein RES-Mittel (Residual Echo Suppression RES, Restechounterdrückung).

30. Verfahren nach einem der Ansprüche 21 bis 29, wobei die Eingangssignale durch einen nichtadaptiven und/oder superdirektiven und/oder MVDR-Strahlformer (Minimum Variance Distortionless Response MVDR, verzerrungsfreie Reaktion mit minimaler Varianz) verarbeitet werden.

31. Computerprogrammerzeugnis, umfassend ein oder mehrere computerlesbare Medien mit computerausführbaren Anweisungen zur Durchführung der Schritte des Verfahrens nach Ansprüchen 21 bis 30.

Revendications

1. Système mains libres pour une utilisation dans un véhicule comprenant un réseau de microphones (101) comportant au moins deux microphones (102), un moyen de traitement du signal et un filtre adaptatif de post traitement (108), dans lequel le moyen de traitement du signal comprend un dispositif de formation de faisceau comportant une entrée reliée aux deux microphones au moins ainsi qu'une sortie reliée à l'entrée du filtre adaptatif de post traitement, le dispositif de formation de faisceau comprenant un moyen de déplacement de faisceau et un moyen de sommation, dans lequel le filtre adaptatif de post traitement est un filtre de Wiener à phase linéaire.

2. Système mains libres selon la revendication 1, dans lequel le filtre adaptatif de post traitement est un filtre dans le

domaine du temps.

- 5 3. Système mains libres selon l'une des revendications précédentes, dans lequel le moyen de traitement du signal comprend en outre au moins deux filtres adaptatifs (110) comportant une entrée reliée à la sortie du moyen de déplacement de faisceau (104) et une sortie reliée au filtre adaptatif de post traitement, les au moins deux filtres adaptatifs étant configurés pour déterminer des paramètres de filtre adaptatif pour le filtre adaptatif de post traitement.
- 10 4. Système mains libres selon la revendication 3, dans lequel, pour chacun des au moins deux microphones, un filtre adaptatif est prévu, lequel possède une entrée reliée à la sortie du moyen de déplacement de faisceau.
- 15 5. Système mains libres selon la revendication 3 ou 4, dans lequel une entrée de chacun des au moins deux filtres adaptatifs est en outre reliée à la sortie du dispositif de formation de faisceau.
- 20 6. Système mains libres selon l'une des revendications précédentes, dans lequel le moyen de traitement du signal comprend en outre un filtre de préaccentuation (111), comprenant en particulier un filtre pour rendre blanche la distribution spectrale d'un signal, comportant une entrée reliée à une sortie du filtre adaptatif de post traitement et/ou d'un filtre de préaccentuation, comprenant en particulier un filtre pour rendre blanche la distribution spectrale d'un signal, comportant une entrée reliée à la sortie du moyen de déplacement de faisceau et une sortie reliée vers les au moins deux filtres adaptatifs.
- 25 7. Système mains libres selon l'une des revendications précédentes, dans lequel le moyen de traitement du signal comprend en outre un filtre inverse (106), en particulier un filtre inverse de distorsion.
- 30 8. Système mains libres selon l'une des revendications précédentes, dans lequel le moyen de traitement du signal comprend en outre un filtre non adaptatif de post traitement (109) comportant une entrée reliée à une sortie du filtre adaptatif de post traitement.
- 35 9. Système mains libres selon l'une des revendications précédentes, dans lequel le moyen de traitement du signal comprend en outre un dispositif adaptatif d'annulation de bruit (ANC) (1101) et/ou un dispositif d'annulation d'écho acoustique (AEC) (1303).
- 40 10. Système mains libres selon l'une des revendications précédentes, dans lequel le dispositif de formation de faisceau est un dispositif non adaptatif de formation de faisceau et/ou un dispositif super directif de formation de faisceau et/ou un dispositif de formation de faisceau à réponse sans distorsion à variance minimale (MVDR).
- 45 11. Système mains libres selon l'une des revendications précédentes, dans lequel le réseau de microphones comprend au moins deux microphones agencés selon une orientation à rayonnement longitudinal par rapport à une première position.
- 50 12. Système mains libres selon la revendication 11, dans lequel le réseau de microphones comprend au moins deux microphones agencés selon une orientation à rayonnement longitudinal par rapport à une seconde position.
- 55 13. Système mains libres selon la revendication 12, dans lequel les au moins deux microphones selon la première orientation à rayonnement longitudinal et les au moins deux microphones selon la seconde orientation à rayonnement longitudinal possèdent un microphone en commun.
14. Système mains libres selon l'une des revendications précédentes, dans lequel le moyen de traitement du signal comprend au moins deux dispositifs de formation de faisceau.
15. Système mains libres selon la revendication 14, comprenant en outre un détecteur d'activité vocale (VAD) et/ou un moyen de commande de commutation (1510).
16. Système mains libres selon l'une des revendications précédentes, comprenant en outre un moyen de suppression d'écho résiduel (RES) et/ou une commande dynamique du volume (DVC) (1510).
17. Système mains libres selon l'une des revendications précédentes, dans lequel le réseau de microphones comprend au moins deux sous réseaux (701, 702, 703).

18. Système mains libres selon l'une des revendications précédentes, comprenant en outre un cadre dans lequel chaque microphone du réseau de microphones est agencé selon une position prédéfinie, en particulier fixe, dans ou sur le cadre.
- 5 19. Système mains libres selon l'une des revendications précédentes, dans lequel au moins un microphone est un microphone directionnel (1001), possédant en particulier une caractéristique cardioïde et/ou étant un microphone différentiel.
- 10 20. Véhicule comprenant un système mains libres conforme à l'une des revendications précédentes.
21. Procédé de réduction de bruit dans un système mains libres de véhicule, comprenant
la réception de signaux d'entrée résultant d'un réseau de microphones comportant au moins deux microphones,
le traitement des signaux d'entrée par un dispositif de formation de faisceau afin de fournir un signal formé en faisceau, et
15 le filtrage adaptatif d'un signal résultant du signal formé en faisceau grâce à un filtre adaptatif de post traitement, dans lequel le filtre adaptatif de post traitement est un filtre de Wiener à phase linéaire.
22. Procédé selon la revendication 21, dans lequel le filtrage adaptatif est exécuté dans le domaine du temps.
- 20 23. Procédé selon la revendication 21 ou 22, comprenant en outre
la fourniture d'au moins deux filtres adaptatifs, en particulier des filtres de Wiener, où
le traitement des signaux d'entrée par un dispositif de formation de faisceau comprend le déplacement de faisceau
des signaux d'entrée afin de fournir des signaux à faisceau déplacé correspondant à l'un des au moins deux
microphones, ainsi que la sommation des signaux, et
25 le filtrage adaptatif comprend la réception et le traitement d'au moins un signal à faisceau déplacé par au moins l'un des au moins deux filtres adaptatifs afin de déterminer des paramètres de filtre adaptatif pour le filtre adaptatif de post traitement.
- 30 24. Procédé selon la revendication 23, dans lequel le filtrage adaptatif comprend en outre la réception d'un signal résultant du signal formé en faisceau par au moins un filtre adaptatif et dans lequel le traitement du signal à faisceau déplacé comprend la détermination de paramètres de filtre adaptatif en utilisant le ou les signaux à faisceau déplacé et le signal résultant du signal formé en faisceau.
- 35 25. Procédé selon la revendication 23 ou 24, dans lequel, pour chaque signal à faisceau déplacé, un filtre adaptatif est prévu pour déterminer des paramètres de filtre adaptatif en utilisant le signal à faisceau déplacé et le signal résultant du signal formé en faisceau.
- 40 26. Procédé selon l'une des revendications 23 à 25, dans lequel la réception d'au moins un signal à faisceau déplacé par au moins l'un des au moins deux filtres adaptatifs comprend le traitement du ou des signaux à faisceau déplacé grâce à un filtre de préaccentuation, comprenant en particulier un filtre pour rendre blanche la distribution spectrale d'un signal.
- 45 27. Procédé selon l'une des revendications 21 à 26, comprenant en outre le traitement d'un signal résultant du réseau de microphones par un filtre inverse, en particulier un filtre inverse de distorsion.
- 50 28. Procédé selon l'une des revendications 21 à 27, comprenant en outre un filtrage non adaptatif d'un signal résultant du signal filtré de manière adaptative et/ou un traitement d'un signal résultant du signal filtré de manière adaptative grâce à un filtre de préaccentuation.
- 55 29. Procédé selon l'une des revendications 21 à 28, comprenant en outre le traitement d'un signal résultant du réseau de microphones, en particulier résultant du signal formé en faisceau, par un dispositif d'annulation adaptatif de bruit (ANC) et/ou par un dispositif d'annulation d'écho acoustique (AEC) et/ou par un moyen de suppression d'écho résiduel (RES).
30. Procédé selon l'une des revendications 21 à 29, dans lequel les signaux d'entrée sont traités par un dispositif de formation de faisceau non adaptatif et/ou super directif et/ou à réponse sans distorsion à variance minimale (MVDR).
31. Produit de programme informatique comprenant un ou plusieurs supports pouvant être lus par ordinateur comportant

EP 1 538 867 B1

des instructions exécutables par ordinateur afin d'exécuter les étapes du procédé conformes aux revendications 21 à 30.

5

10

15

20

25

30

35

40

45

50

55

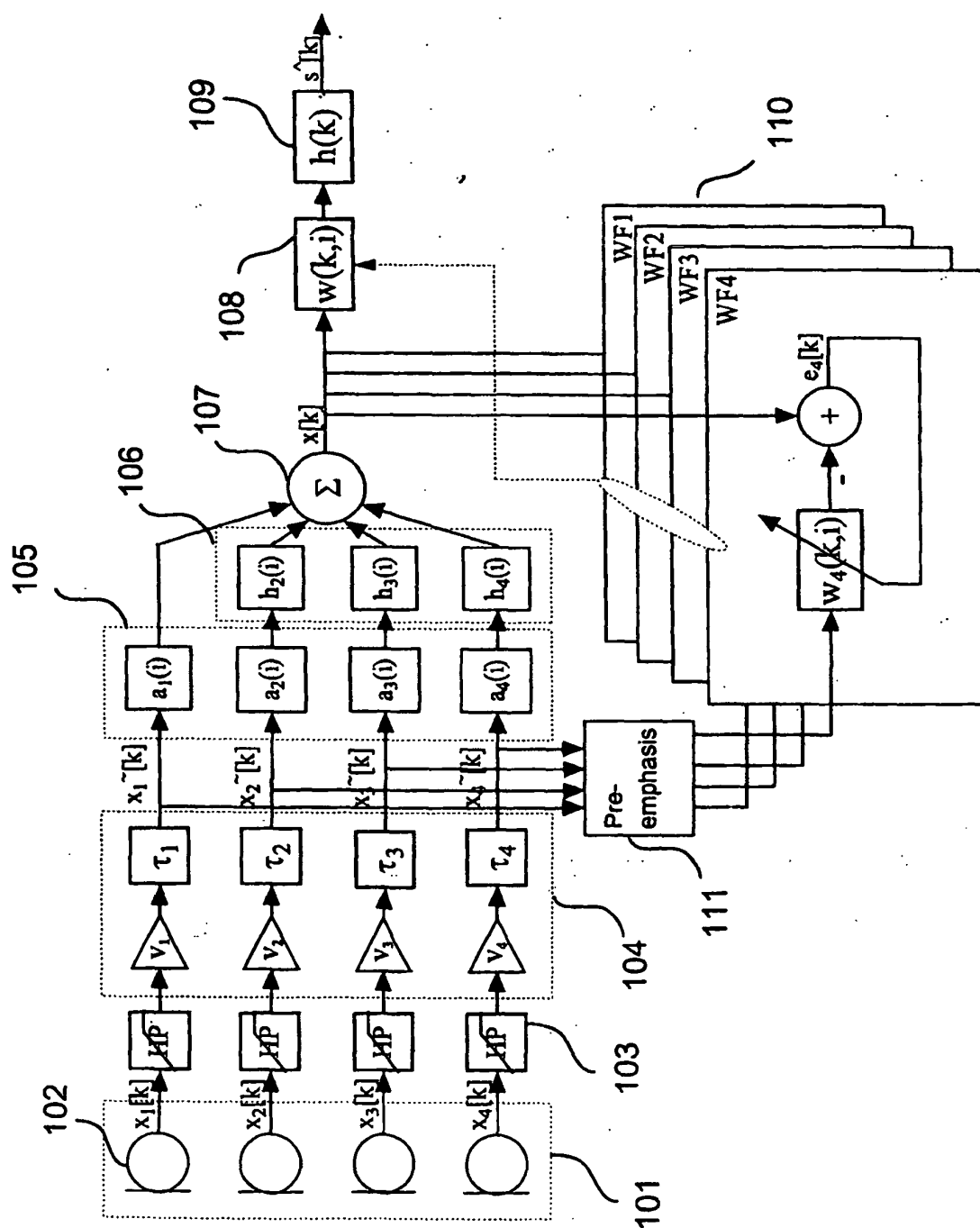


Fig. 1

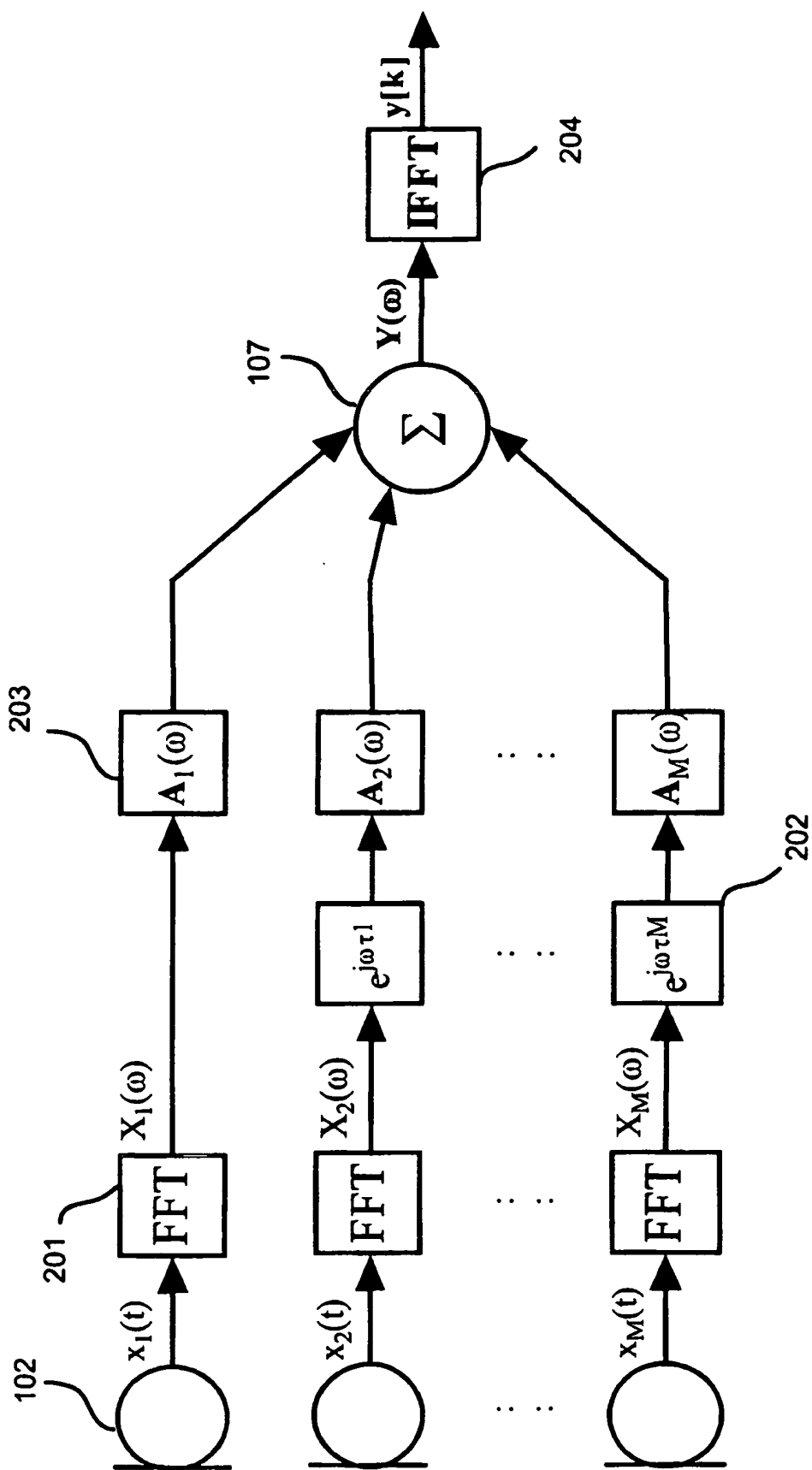


Fig. 2

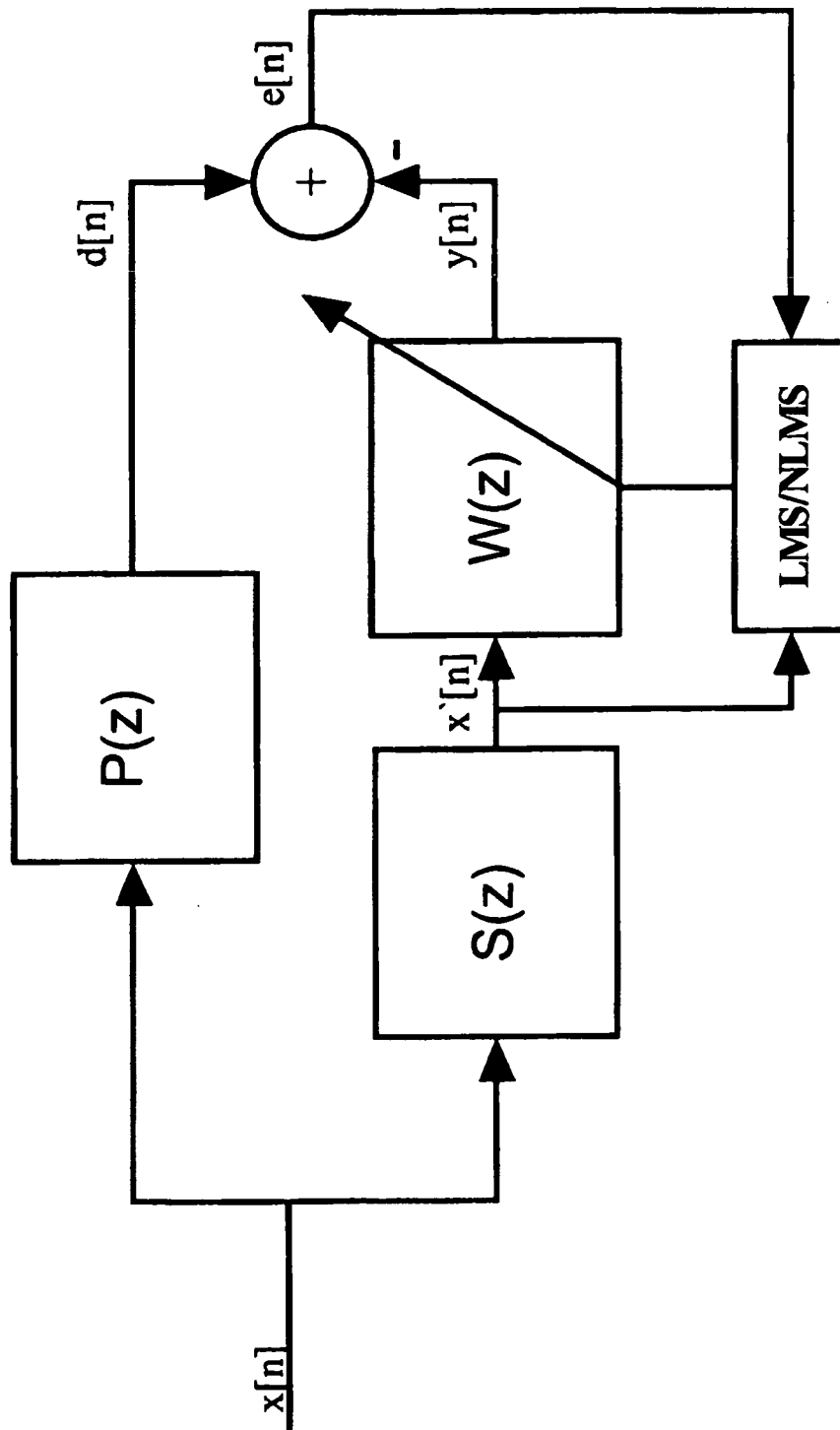


Fig. 3

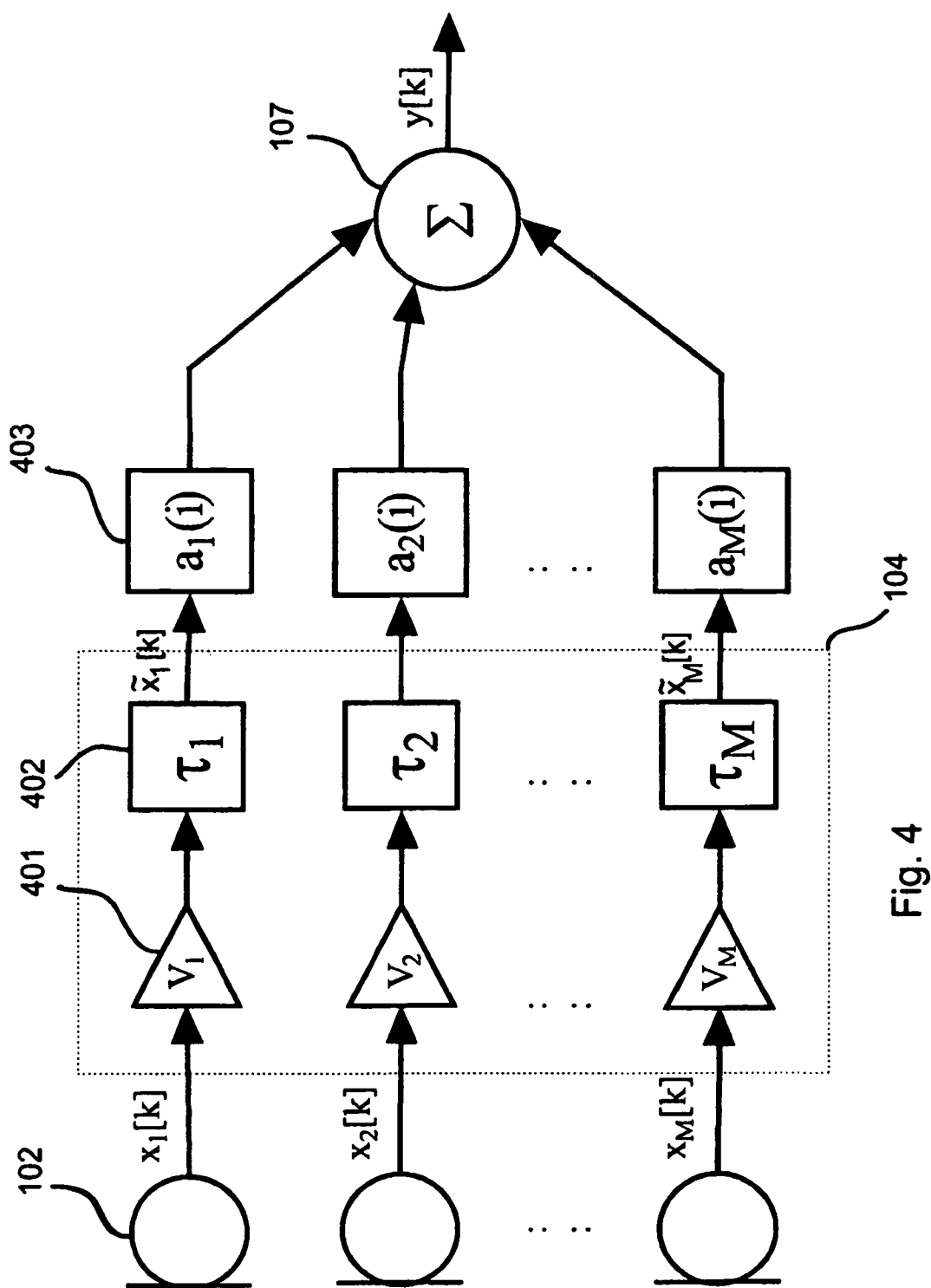


Fig. 4

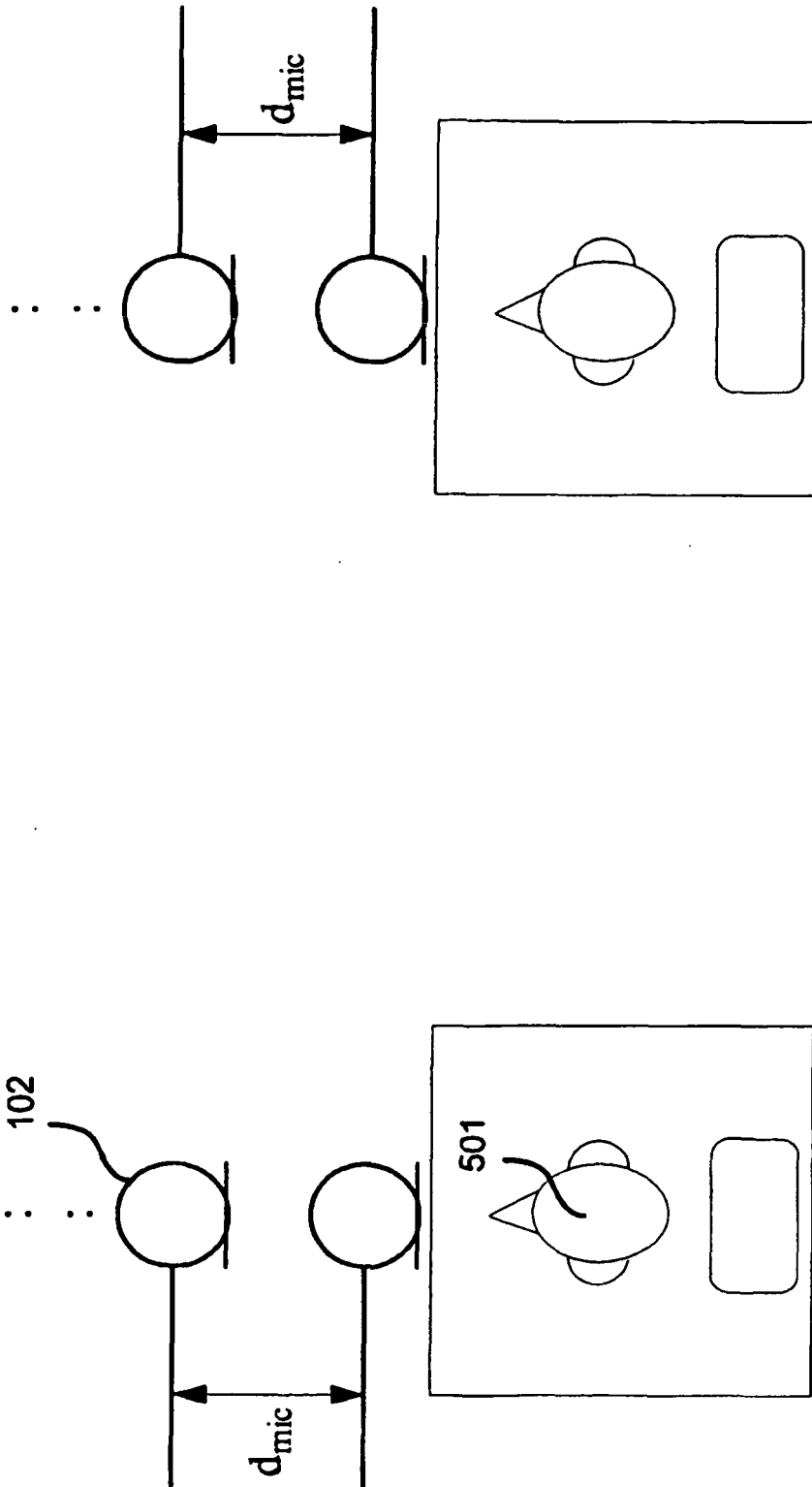


Fig. 5A

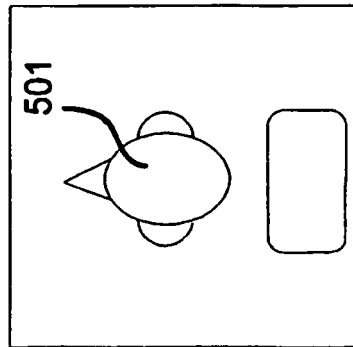
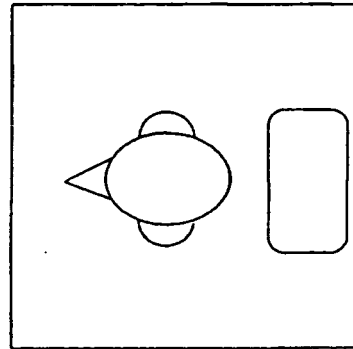
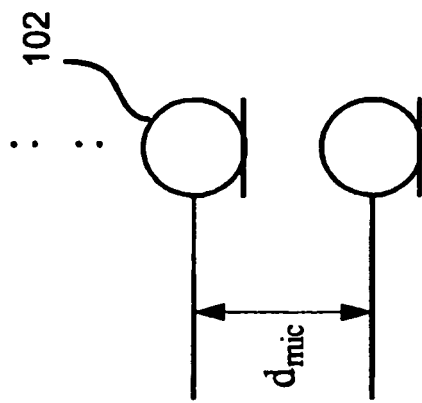


Fig. 5B

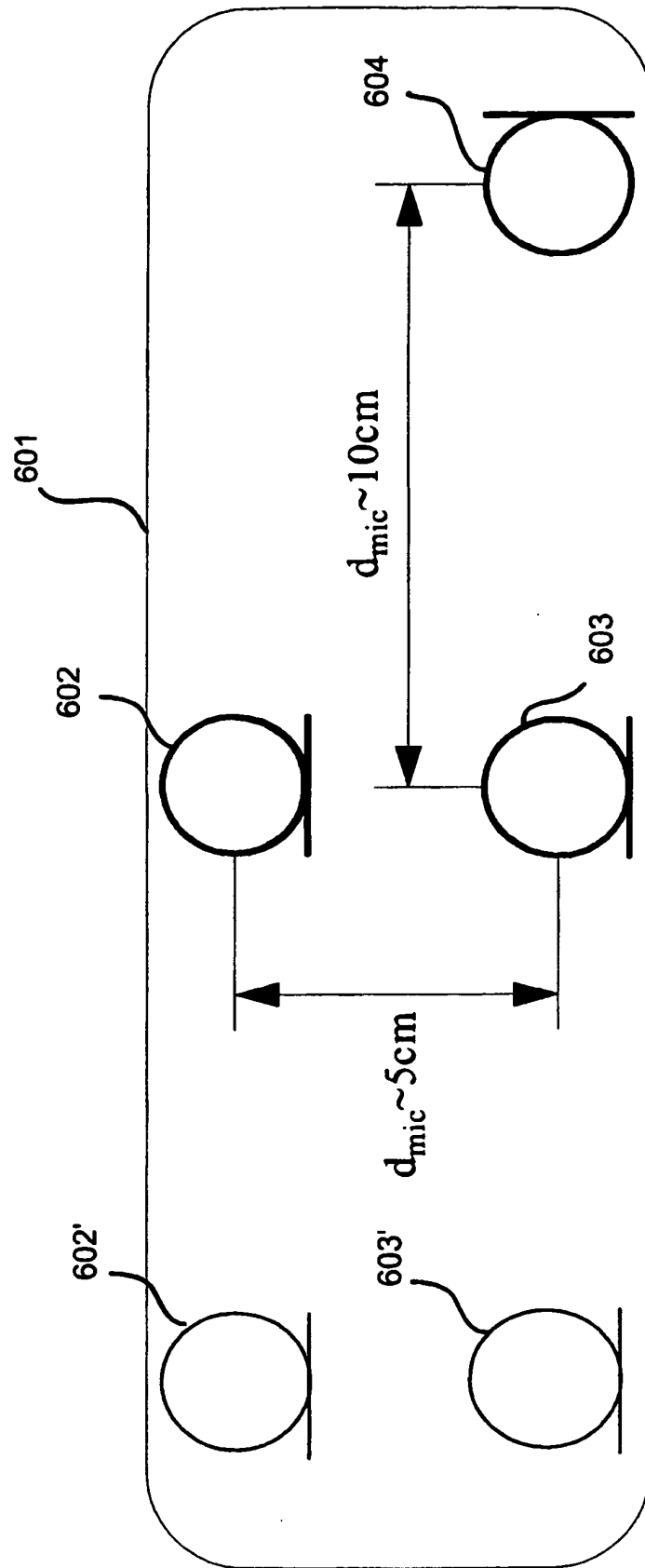


Fig. 6A

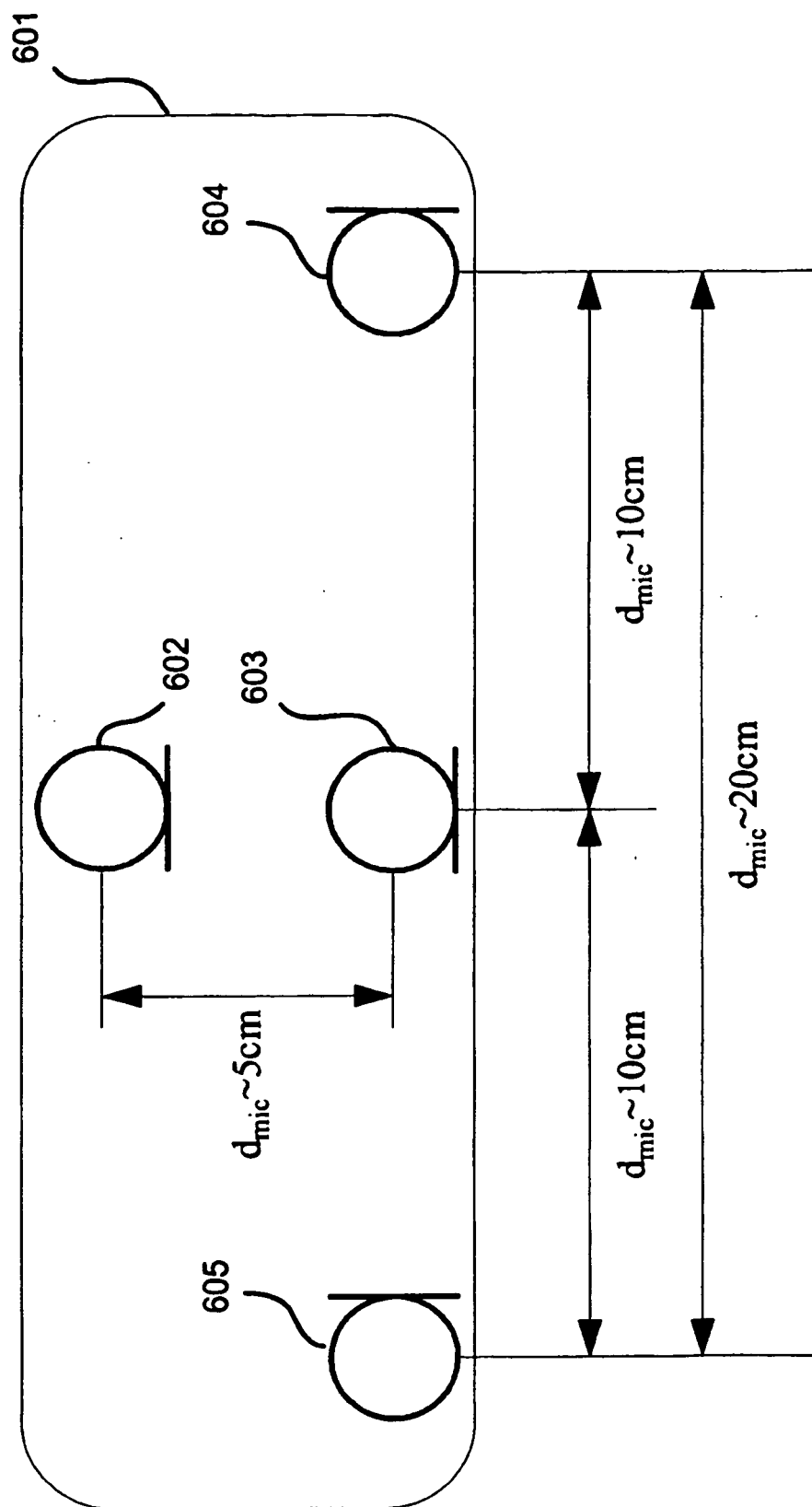


Fig. 6B

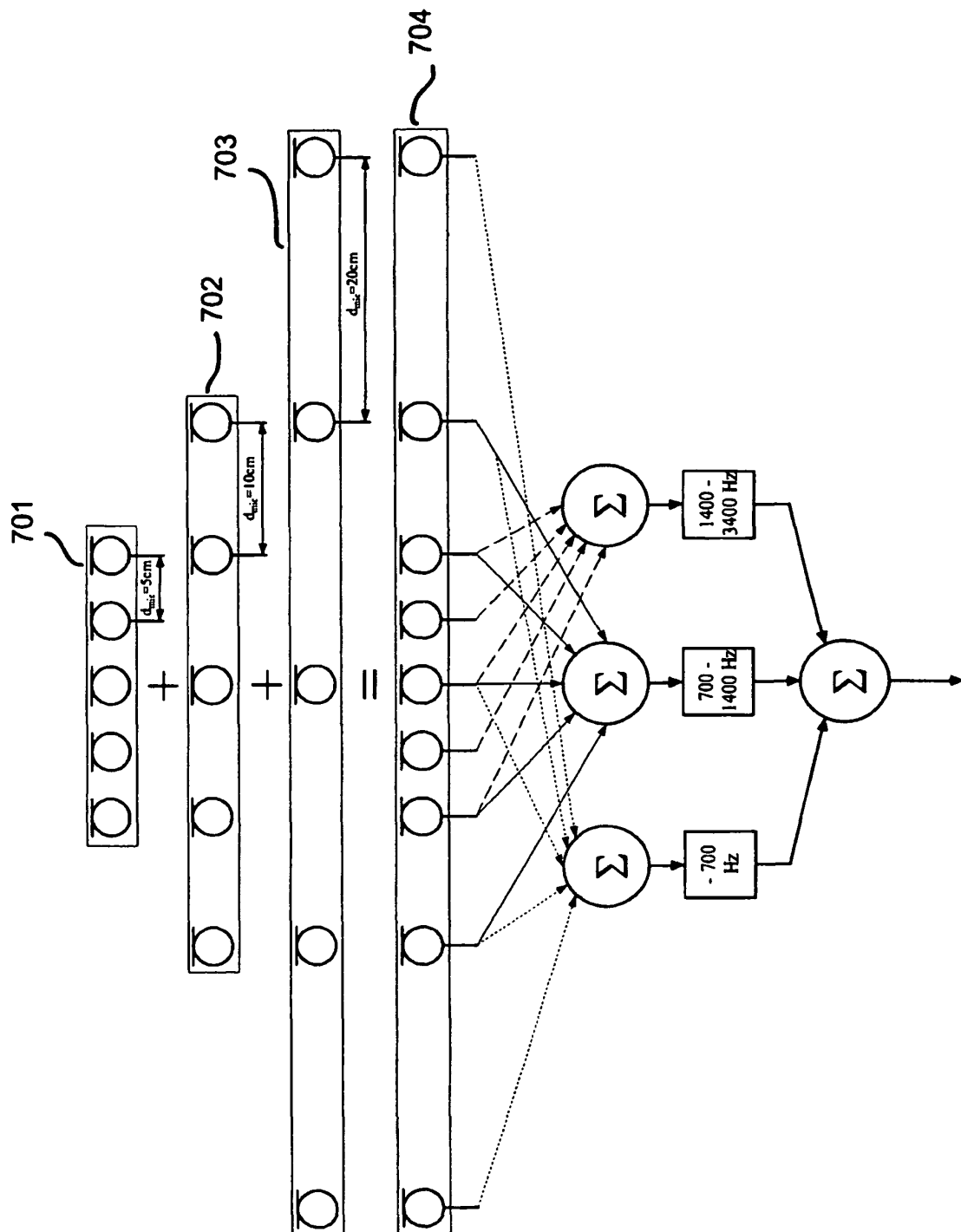


Fig. 7

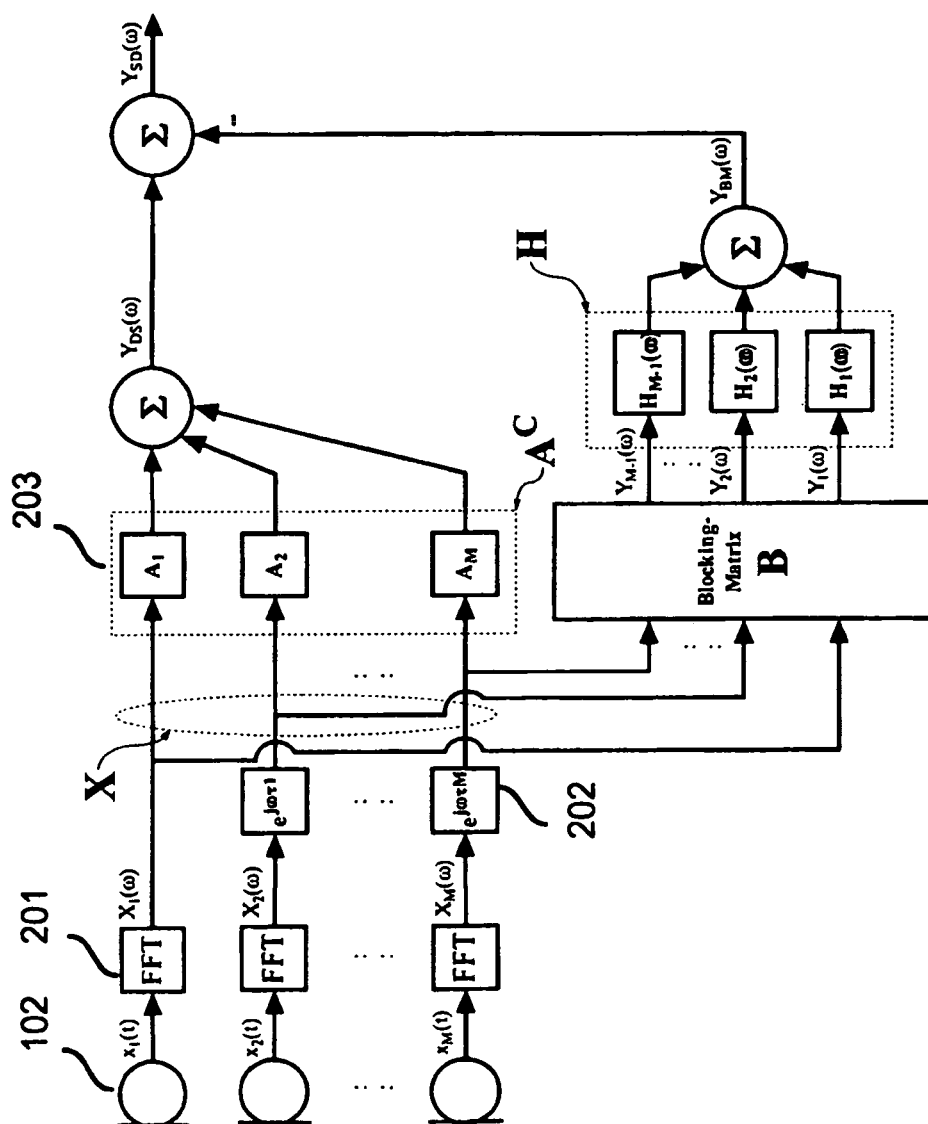


Fig. 8

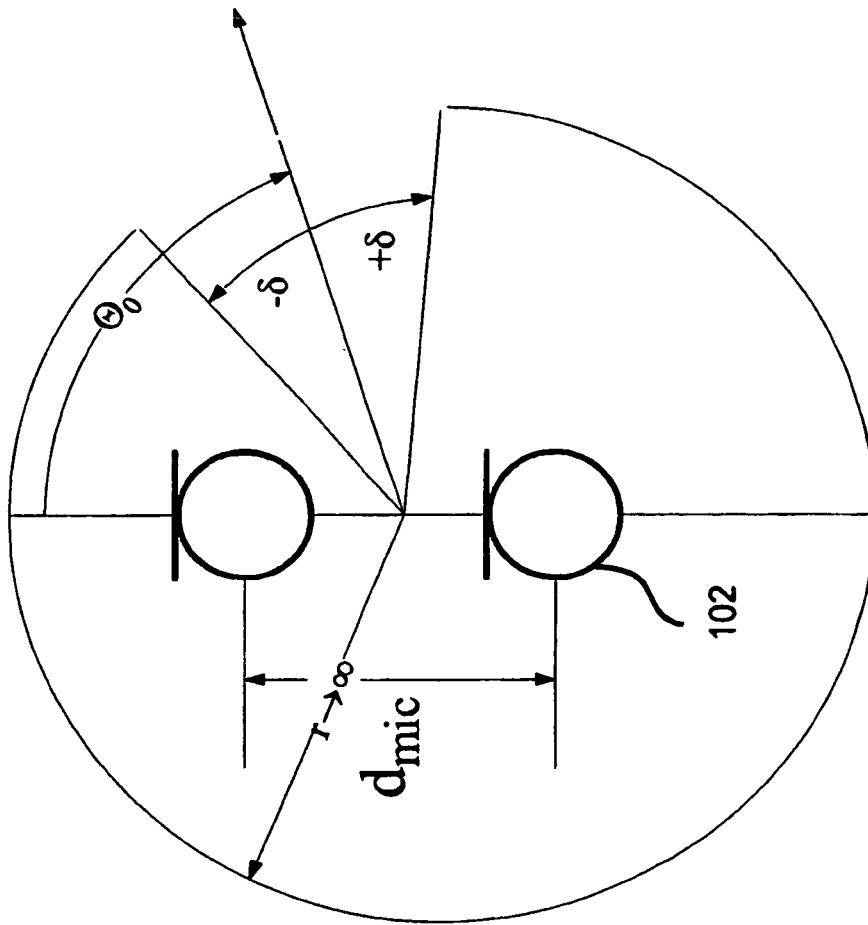


Fig. 9

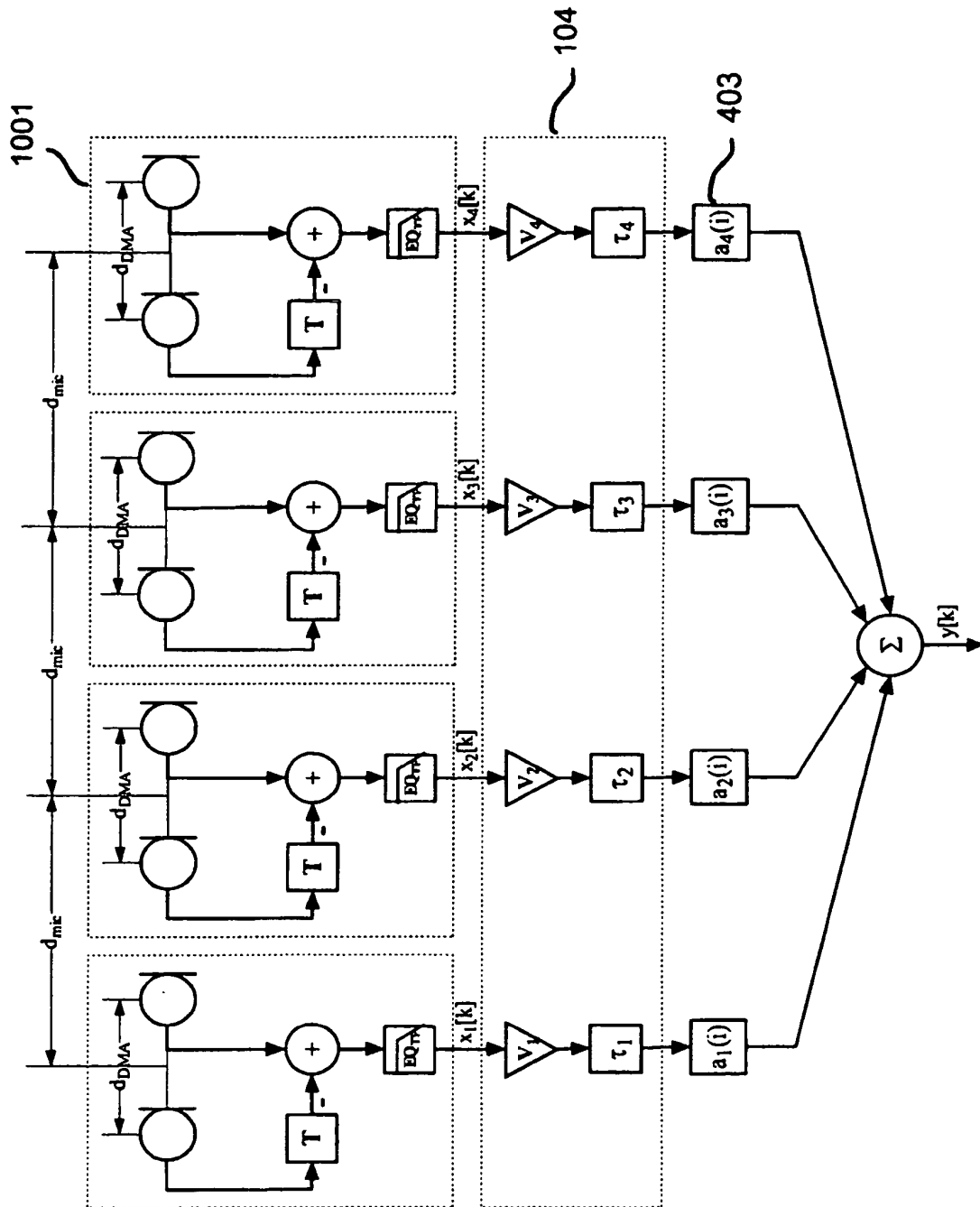


Fig. 10

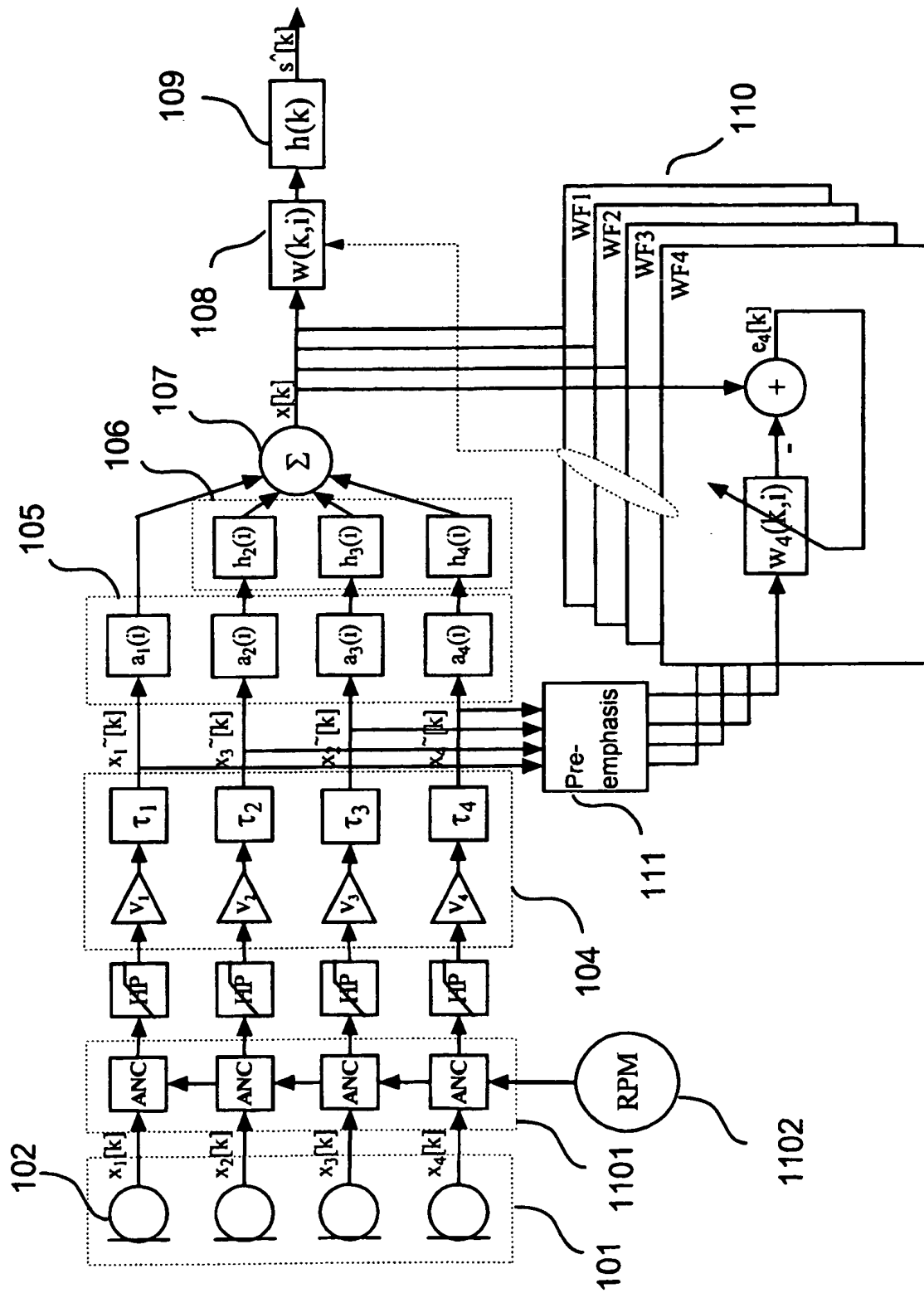


Fig. 11

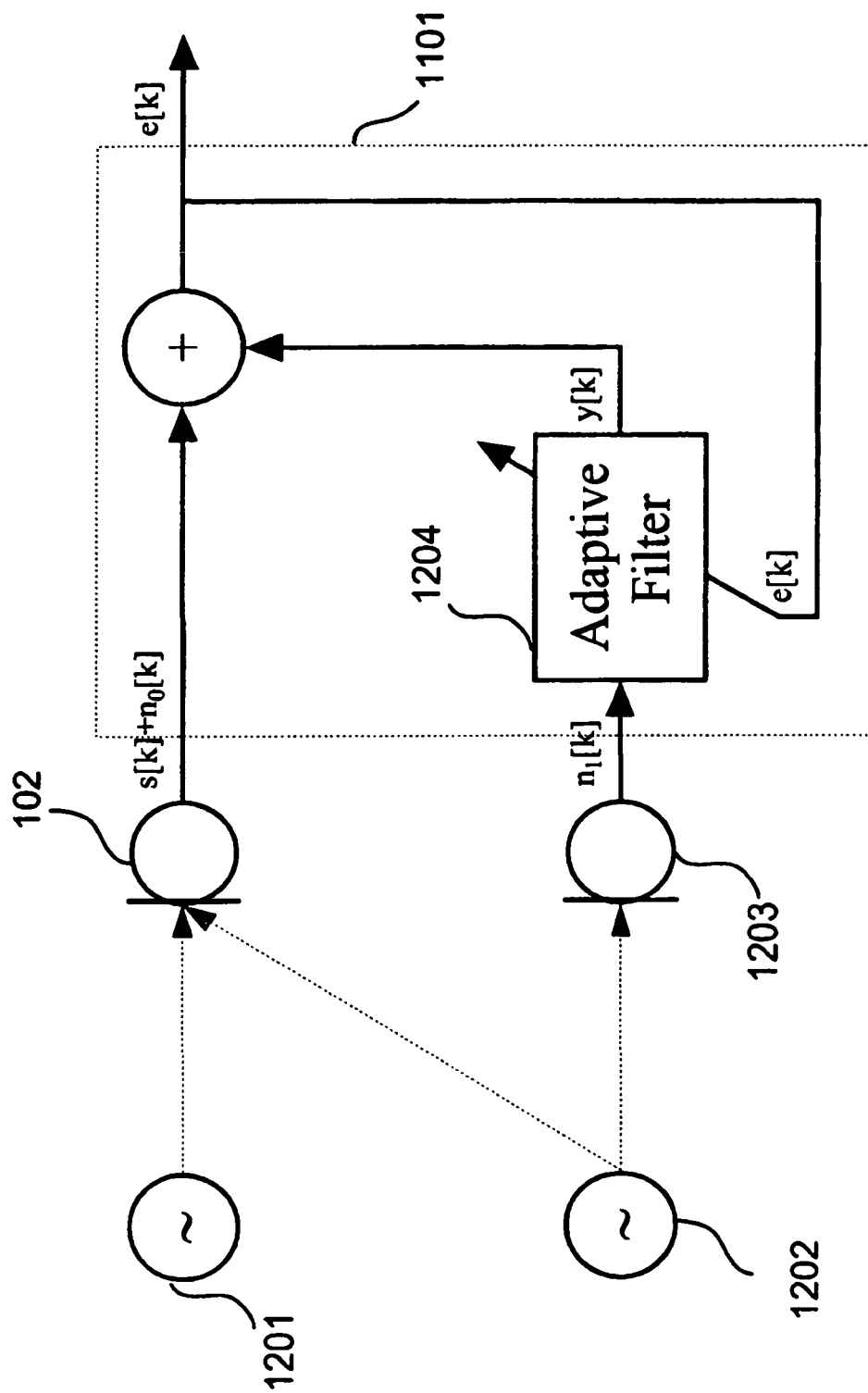


Fig. 12

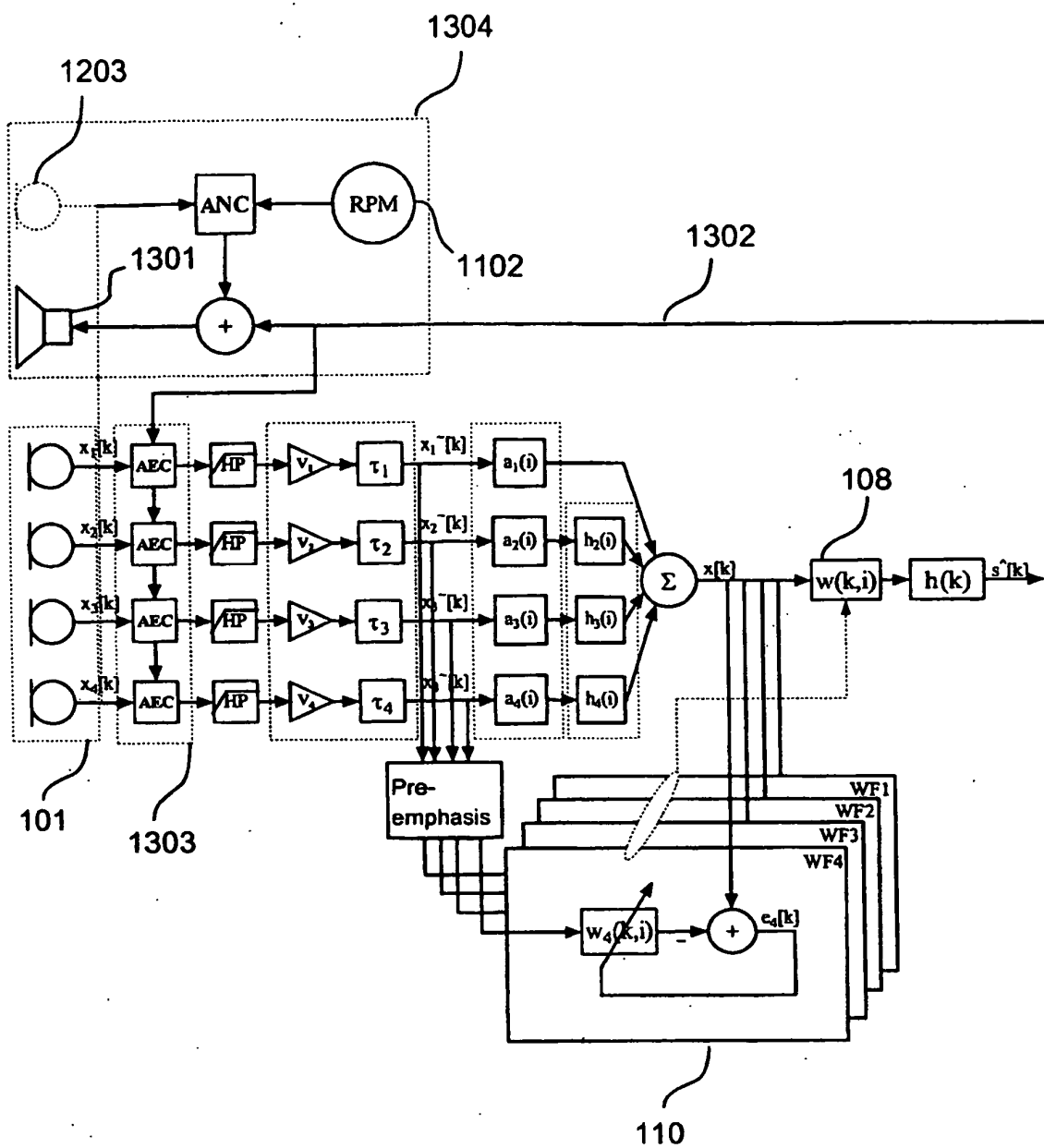


Fig. 13

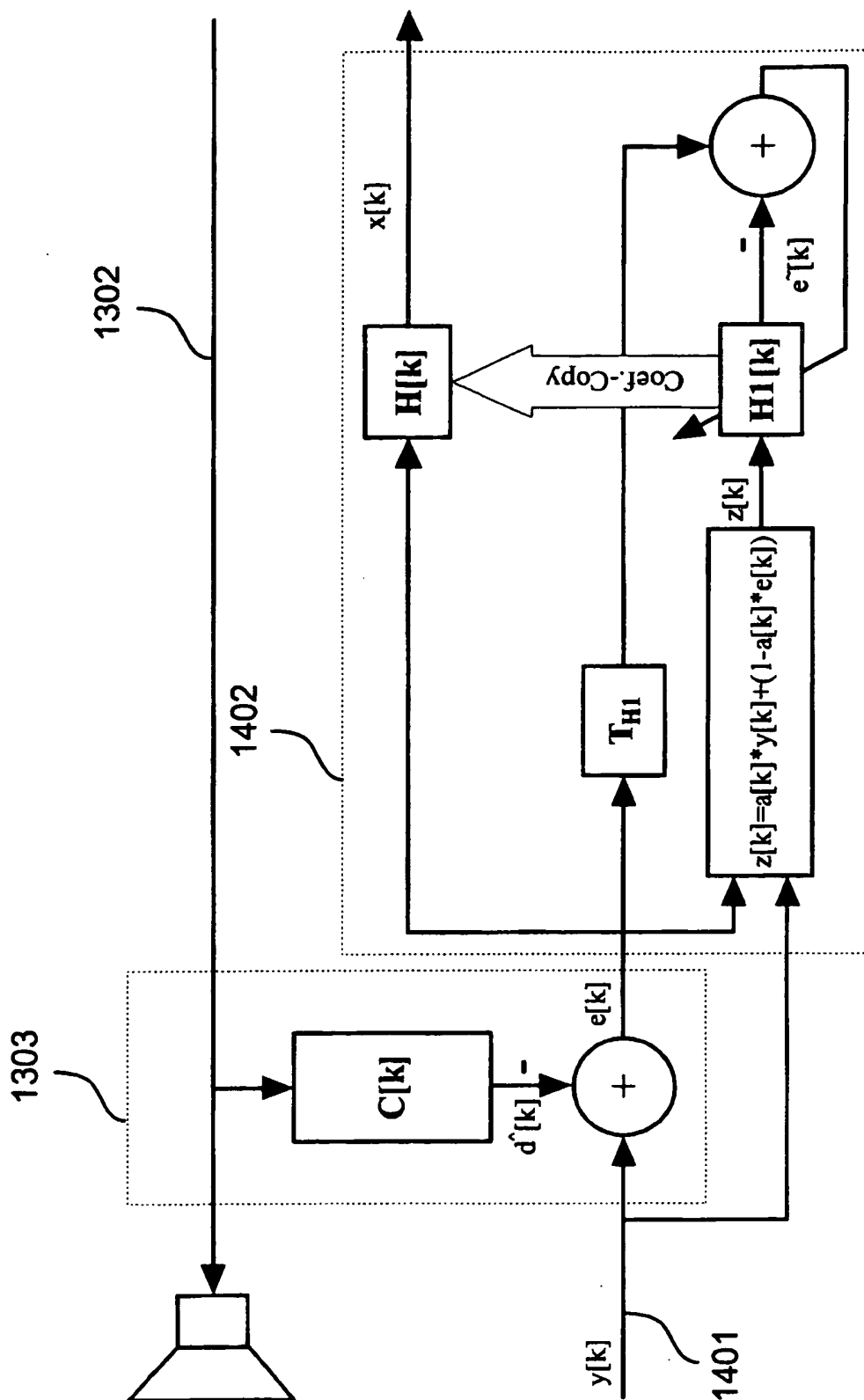


Fig. 14

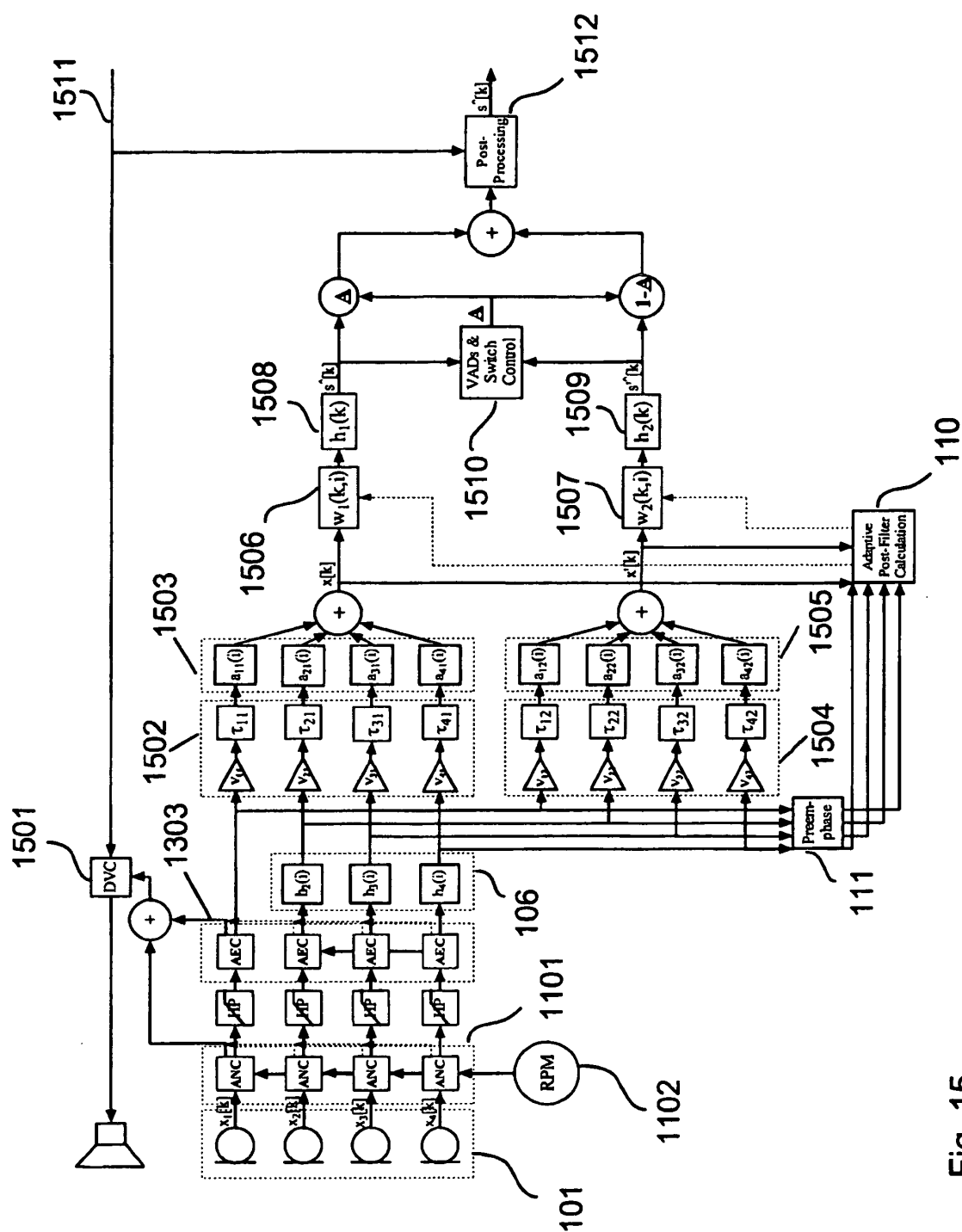


Fig. 15

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

- WO 03015464 A [0003]