(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 March 2016 (31.03.2016)

WIPO | PCT

(10) International Publication Number
WO 2016/045706 A1

(51) International Patent Classification: *H04R 3/00* (2006.01)

(21) International Application Number: PCT/EP2014/070243

(22) International Filing Date: 23 September 2014 (23.09.2014)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant: **BINAURIC SE** [DE/DE]; Am Söldnermoos 17, 85399 Hallbergmoos (DE).

(72) Inventors: **KRÜGER, Hauke**; Gallierstr. 33, 52074 Aachen (DE). **GEISER, Bernd**; Weidenweg 10, 52074 Aachen (DE).

(74) Agent: **EISENFÜHR SPEISER PATENTANWÄLTE RECHTSANWÄLTE PARTGMBB**; P.O. Box 31 02 60, 80102 Munich (DE).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) Title: METHOD AND APPARATUS FOR GENERATING A DIRECTIONAL SOUND SIGNAL FROM FIRST AND SECOND SOUND SIGNALS

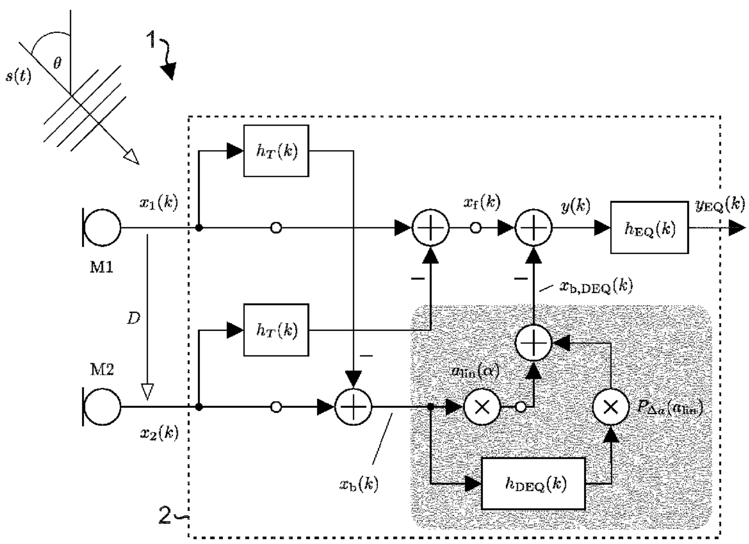


FIG. 4

(57) Abstract: The present invention relates to a method for generating a directional sound signal *($y(k)$)* from first and second sound signals *($x_1(k)$, $x_2(k)$)*, which are generated by a first and a second microphone (M1, M2), which are separated by a distance (D). The method comprises generating first and second differential sound signals *($x_f(k)$, $X_{b,DEQ}(k)$)* based on the first and second sound signals *($x_1(k)$, $x_2(k)$)*, and generating the directional sound signal *($y(k)$)* according to a frequency-dependent directional response pattern based on the first and second differential sound signals *($x_f(k)$, $X_{b,DEQ}(k)$)*. The generating of the second differential sound signal ($X_{b,DEQ}(k)$) comprises generating a difference signal *($xb(k)$)* of the first and the second sound signals *($x_1(k)$, $x_2(k)$)* and a frequency-selective processing that depends on a steering angle (a), which indicates a desired direction of maximum attenuation of the frequency-dependent directional response pattern, wherein the frequency-selective processing adjusts the actual direction of maximum attenuation of the frequency-dependent directional response pattern to correspond to the steering angle (α) substantially independent of frequency (ω) over the frequency range of the directional sound signal *($y(k)$)*.

WO 2016/045706 A1

Method and apparatus for generating a directional sound signal from first and second sound signals

FIELD OF THE INVENTION

The present invention generally relates to the field of sound acquisition. More particularly, the present invention relates to a method and an apparatus for generating a directional sound signal from first and second sound signals, which are generated by a first and a second microphone, which are separated by a distance.

BACKGROUND OF THE INVENTION

1. Introduction

For sound acquisition in realistic acoustic environments, microphone arrays proved to be useful. They are designed to attenuate possible noise and interference components while retaining the desired source signal by exploiting different spatial (or directional) characteristics of the different signal sources (see, e.g., J. Benesty, J. Chen, and Y. Huang, “Microphone Array Signal Processing,” Heidelberg: Springer, 2008 for an overview).

A simple, yet efficient approach is the first-order differential microphone array described in G. Elko and A.-T. N. Pong, “A simple adaptive first-order differential microphone,” in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), pages 169 to 172, October 1995. This microphone array, which is schematically and exemplarily shown in Fig. 1 allows to place two symmetrical notches (directions of maximum attenuation) at angles of θ and $360° - \theta$. Herein, three independent enhancements to the original method are proposed and a practical implementation for handsfree communication is described. A possible target device is a wireless loudspeaker with two integrated miniature digital micro-electromechanical system (MEMS) microphone capsules which facilitate handsfree audio communication.

2. Differential Microphone Array

First, the originally proposed method shall be reviewed with reference to Fig. 1, which shows schematically and exemplarily a differential microphone array according to G. Elko and A.-T. N. Pong. Two closely spaced omnidirectional microphones M1 and M2 are used to capture the acoustic environment. The corresponding digital signals $x_1(k)$ and $x_2(k)$ are sampled with a rate of f_s. Due to the small distance D between M1 and M2, a coherent mutual subtraction – for convenience, acausal filters are assumed herein; in practice, appropriate signal alignment is required as marked by ‘o’ in Figs. 1 and 4 – of the aligned signals, i.e.,

$$x_f(k) = x_1(k) - x_2(k) \star h_T(k) \tag{1}$$

$$x_b(k) = x_2(k) - x_1(k) \star h_T(k) \tag{2}$$

can be achieved. Thereby, $h_T(k)$ is a fractional delay filter with a delay time of $T = D/c$ (c: speed of sound). This value corresponds to the sound propagation time from one microphone to the other. The signals $x_f(k)$ and $x_b(k)$ can be interpreted as “forward and backward facing cardioid” signals as the respective directional responses of Eqs. (1) and (2) form cardioid shapes (see Fig. 3 in G. Elko and A.-T. N. Pong). For example $x_f(k)$ does not contain any sound components from the rear direction ($\theta = 180°$) while $x_b(k)$ does not contain components from the front direction ($\theta = 0°$). Both signals $x_f(k)$ and $x_b(k)$ are finally combined according to

$$y_{EQ}(k) = h_{EQ}(k) \star (x_f(k) - a \cdot x_b(k)) \tag{3}$$

whereby the low-pass equalizer $h_{\mathrm{EQ}}(k)$ compensates for the highpass effect of the coherent subtraction operations and the scalar factor a can be used to control the notch angle(s): $a = 0$ corresponds to a (double) notch angle of $180°$ while $a = 1$ corresponds to two symmetrical notch angles of $90°$ and $270°$. Due to the symmetry, all angle specifications are restricted to $0° \ldots 180°$ in the following, whereby any statement for a specific angle θ also applies to its symmetrical counterpart $360° - \theta$. In the common operation scenario of the microphone array, the desired sound source lies in the front half plane ($\theta = 0° \ldots 90°$) while the undesired noise or interference source(s) lie(s) in the rear half plane ($\theta = 90° \ldots 180°$).

A common problem with differential microphone arrays are the tolerances of the employed microphones, leading to a “microphone mismatch” and therefore noise amplification (see M. Buck and M. Rößler, “First Order Differential Microphone Arrays for Automotive Applications,” in Proceedings of International Workshop on Acoustic Echo and Noise Control (IWAENC), September 2001). The digital MEMS microphones in the possible target device usually exhibit relatively constant frequency responses; therefore, individual microphone equalization is preferably not necessary for the envisaged application. However, their power levels may still vary to a certain extent due to mounting and assembly tolerances, which is disadvantageous since it is preferred to have fully matched input levels in order to utilize the full potential of the method.

SUMMARY OF THE INVENTION

3. Objects and Solutions

It is an object of the present invention to provide a method and an apparatus for generating a directional sound signal from first and second sound signals, which are generated by a first and a second microphone, which are separated by a distance, which inter alia allow for a wider operating frequency and help to reduce a noise gain.

In a first aspect of the present invention, a method for generating a directional sound signal from first and second sound signals, which are generated by a first and a second microphone, which are separated by a distance, is presented, wherein the method comprises:

- generating first and second differential sound signals based on the first and second sound signals, and

- generating the directional sound signal according to a frequency-dependent directional response pattern based on the first and second differential sound signals,

wherein the generating of the second differential sound signal comprises generating a difference signal of the first and the second sound signals and a frequency-selective processing that depends on a steering angle, which indicates a desired direction of maximum attenuation of the frequency-dependent directional response pattern, wherein the frequency-selective processing adjusts the actual direction of maximum attenuation of the frequency-dependent directional response pattern to correspond to the steering angle substantially independent of frequency over the frequency range of the directional sound signal.

The present invention is based on the idea that by employing these steps, a (substantially) frequency invariant notch characteristic can be obtained even for larger microphone distances. A larger distance also helps to confine the noise gain of the array. Therefore, the array becomes practically usable even for higher sampling rates (e.g., $16\mathrm{kHz}$).

The term “difference signal” as used herein, also includes the case where one or both of the first and the second sound signals is/are further temporally delayed, for example, by means of a fractional delay filter $h_T(k)$, as described in section 2 above.

It is preferred that the frequency-selective processing comprises weighting the difference signal with an approximated steering factor that is independent of frequency to generate a weighted difference signal and correcting for the approximation by adding a correction signal that is generated from the difference signal in dependence of frequency and the steering angle.

It is further preferred that the generation of the correction signal comprises applying two separate operations, one being dependent on frequency and independent of the steering angle and one being dependent on the steering angle but independent of frequency.

It is preferred that the generation of the correction signal comprises filtering the difference signal with a filter that is dependent on frequency and independent of the steering angle to generate a filtered difference signal.

It is further preferred that the generation of the correction signal further comprises weighting the filtered difference signal with a factor that is dependent on the steering angle and independent of frequency.

It is preferred that the factor is determined by using a polynomial approximation that is evaluated with the steering angle.

It is further preferred that the method further comprises filtering the directional sound signal with a low-pass filter to generate a filtered directional sound signal.

It is preferred that the approximated steering factor for a time instance is adapted for the following time instance by adding an adaptation value that is scaled by a stepsize parameter, wherein the stepsize parameter is adapted in dependence of estimated energies of coherent and incoherent sound components.

Employing these steps provides a fast notch adaptation algorithm, with which the noise (or interferer) suppression works more reliable in a broader range of acoustic scenarios. Also, the desired source does not compromise the notch adaptation anymore. Moreover, the combination with the directional equalizer leads to a more stable direction of arrival tracking.

It is further preferred that the energy of the incoherent sound components is approximated by the estimated short-term energy of the directional sound signal and the energy of the coherent sound components is approximated by a fraction of the estimated short-term energy of the difference signal.

It is preferred that the method further comprises estimating a relative gain of the first and the second microphone and equalizing power levels of the first and the second microphone based on the relative gain.

Employing these steps enables the array to work reliably despite certain assembly and sensor tolerances.

It is further preferred that the relative gain is determined based on recursively estimated variances of the first and second sound signals.

It is preferred that the first and the second microphone are omnidirectional microphones.

In a second aspect of the present invention, an apparatus for generating a directional sound signal from first and second sound signals, which are generated by a first and a second microphone, which are separated by a distance, is presented, wherein the apparatus comprises:

- first generating means for generating first and second differential sound signals based on the first and second sound signals, and
- second generating means for generating the directional sound signal according to a directional response pattern based on the first and second differential sound signals,

wherein the generating of the second differential sound signal comprises generating a difference signal of the first and the second sound signals and a frequency-selective processing that depends on a steering angle, which indicates a desired direction of maximum attenuation of the frequency-dependent directional response pattern, wherein the frequency-selective processing adjusts the actual direction of maximum attenuation of the frequency-dependent directional response pattern to correspond to the steering angle substantially independent of frequency over the frequency range of the directional sound signal.

In a third aspect of the present invention, a system is presented, wherein the system comprises:

- a first and a second microphone, which are separated by a distance and generate first and second sound signals, and
- an apparatus as defined in claim 13.

In a fourth aspect of the present invention, computer program comprising program code means, which, when run on a computer controlling the apparatus according to claim 13, perform the steps of the method according to any of claims 1 to 12 is presented.

It shall be understood that the method of claim 1, the apparatus of claim 13, the system of claims 14, and the computer program of claim 15 have similar and/or identical preferred embodiments, in particular, as defined in the dependent claims.

It shall be understood that a preferred embodiment of the invention can also be any combination of the dependent claims with the respective independent claim.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereinafter. In the following drawings:

Fig. 1 shows schematically and exemplarily a differential microphone array according to G. Elko and A.-T. N. Pong,

Fig. 2 illustrates the optimal steering factor vs. a linear approximation,

Fig. 3 shows directional frequency responses (rear half plane) of a wideband ($f_\mathrm{s} = 16\,\mathrm{kHz}$) array with $D = 1.8\,\mathrm{cm}$ steered towards $a = 135°$ using the linear approximation of Eq. (7) and the directional equalizer of Fig. 4,

Fig. 4 shows schematically and exemplarily a differential microphone array with directional equalizer according to an embodiment of the present invention,

Fig. 5 illustrates schematically and exemplarily the performance of the proposed fast notch adaptation algorithm ($D = 1.8\,\mathrm{cm},\ f_\mathrm{s} = 16\,\mathrm{kHz}$). The optimal values of a_lin for the incidence angles used herein are: $a_\mathrm{lin}(90°) = 1$, $a_\mathrm{lin}(135°) \approx 0.17$, $a_\mathrm{lin}(180°) = 0$, $a_\mathrm{lin}(0°) = \mathrm{undef}$, and

Fig. 6 shows schematically and exemplarily a developer interface for the exemplarily implemented array based on the RTProc rapid real time prototyping framework (see H. Krüger and P. Vary, “RTPROC: A System for Rapid Real-Time Prototyping in Audio Signal Processing,” in Proceedings of

IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications, October 2008).

DETAILED DESCRIPTION OF EMBODIMENTS

4. Input Level Alignment

In order to fully match the input levels of the microphones M1 and M2, an input level alignment procedure is devised herein.

First, the relative gain $g_\mathrm{c} = \sqrt{\hat{\sigma}_{x_1}^2 / \hat{\sigma}_{x_2}^2}$ is computed on the basis of the recursively estimated input channel variances $\hat{\sigma}_{x_1}^2$ and $\hat{\sigma}_{x_2}^2$. Then, correction gains are obtained

$$g_1 = \begin{cases} g_\mathrm{c}^{-1} & \text{if } \mathrm{g_c} > 1 \\ 1 & \text{else} \end{cases} \quad \text{and} \quad g_2 = \begin{cases} g_\mathrm{c} & \text{if } \mathrm{g_c} < 1 \\ 1 & \text{else} \end{cases} \tag{4}$$

which are finally recursively smoothed before they are applied to the respective input channels. With these gains, the “louder” channel’s level is reduced to the “quieter” channel’s level.

5. Directional Equalization

In the following, it will become clear that the directional response of the original method described in G. Elko and A.-T. N. Pong is distorted, i.e., the actual notch angle clearly deviates from the desired angle, in particular for high frequencies. To correct this behavior, we herein propose an efficient “directional equalization” approach.

It can be shown (proof omitted) that the angle-dependent transfer function of the differential array (without the output equalizer) is

$$H(\omega, \phi) = 2j \cdot \mathrm{e}^{-j\frac{\omega D}{2\mathrm{c}}(\cos\phi + 1)} \cdot \left[\sin\left(\frac{\omega D}{2\mathrm{c}}(1 + \cos\phi)\right) - a \cdot \sin\left(\frac{\omega D}{2\mathrm{c}}(1 - \cos\phi)\right)\right] \tag{5}$$

with the (angular) frequency $\omega = \Omega \cdot f_\mathrm{s} = 2\pi f$ and the respective angle of observation ϕ. Now this transfer function should become zero for a *specific* angle, i.e., the so-called *steering angle* α. Hence, by requiring $H(\omega, \phi = \alpha) = 0$, the optimal *steering factor* a can be deduced from Eq. (5):

$$a(\omega, \alpha) = \frac{\sin\left(\frac{\omega D}{2c}(1 + \cos\alpha)\right)}{\sin\left(\frac{\omega D}{2c}(1 - \cos\alpha)\right)} \quad (6)$$

which obviously depends on α and on the frequency ω. The steering angle α should, ideally, be adapted to match the interference or noise incidence angle θ. This is discussed in section 6.

We note that if we approach the spatial alias frequency, i.e., $\omega \rightarrow \omega_\mathrm{A} = \frac{\pi c}{D}$, we have $a(\omega, \alpha) \rightarrow 1$ regardless of the angle α, i.e., the steering angle is not controllable at ω_A, which renders it a natural cutoff frequency for the entire array. The frequency dependency of a is, however, a contradiction to the scalar multiplication operation in Eq. (3). Therefore, usually, small values of ωD are assumed and a linear approximation of Eq. (6) is used (see H. Puder, “Acoustic noise control: An overview of several methods based on applications in hearing aids,” in IEEE Pacific Rim Conference on Comm. Computers and Signal Processing, pages. 871 to 876, August 2009):

$$a_{\mathrm{lin}}(\alpha) = \frac{1 + \cos\alpha}{1 - \cos\alpha} \quad (7)$$

This approximation is frequency independent which is sufficient for many applications. However, $a_{\mathrm{lin}}(\alpha)$ deviates more and more from the optimal $a(\omega, \alpha)$ as ω approaches ω_A, which is illustrated in Fig. 2. The effect already becomes relevant for frequencies well below ω_A and also varies with the angle α. For a clear demonstration, $\alpha = 135°$ will be used in the following examples, because the deviation from the expected behavior is very prominent in this case. The directional response of a differential microphone array using the linear approximation according to Eq. (7) is shown in Fig. 3 (a).

To correct the obvious deformation of the notch characteristic, we propose to integrate the optimal (frequency and angle dependent) steering factor $a(\omega, \alpha)$ into the time domain realization of the differential array. In principle, this can be achieved by replacing the scalar multiplication operation (factor a in Eq. (3) and in Fig. 1) with a filtering operation, whereby the angle dependent filter transfer functions are given by $a(\omega, \alpha)$. For a more efficient implementation with low memory and computational requirements, we consider the approximation error of the steering factor:

$$\Delta a(\omega, \alpha) = a(\omega, \alpha) - a_{\text{lin}}(\alpha) \tag{8}$$

Fortunately, $\Delta a(\omega, \alpha)$ is separable with good accuracy:

$$\Delta a(\omega, \alpha) \approx \Delta a(\omega) \cdot \Delta a(\alpha) \tag{9}$$

The factors $\Delta a(\omega)$ and $\Delta a(\alpha)$ can be computed by marginalization of the 2-dimensional function $\Delta a(\omega, \alpha)$ and appropriate normalization. The factor $\Delta a(\omega)$ can now be regarded as the frequency response of a *fixed* filter. It can be transformed to the time domain via periodic extension, inverse DFT, cyclic shifting (to enforce causality) and an appropriate shortening to a desired length. The resulting FIR filter coefficients $h_{\text{DEQ}}(k)$, e.g., of order 16, are independent of the steering angle α. The angular dependency is then reintroduced with a polynomial approximation (e.g., order 4) of the second factor $\Delta a(\alpha)$ after a variable transformation from α to $a_{\text{lin}}(\alpha)$, i.e., $P_{\Delta a}(a_{\text{lin}}(\alpha)) \approx \Delta a(\alpha(a_{\text{lin}}))$.

The resulting differential microphone array with directional equalization is shown schematically and exemplarily in Fig. 4, whereby the “directionally equalized” backward cardioid signal $x_{\text{b,DEQ}}(k)$ is given as

$$\begin{aligned} x_{\text{b,DEQ}}(k) &= \mathcal{F}^{-1}\{a(\omega, \alpha)\} * x_{\text{b}}(k) \\ &\approx \mathcal{F}^{-1}\{a_{\text{lin}}(\alpha) + \Delta a(\omega) \cdot \Delta a(\alpha)\} * x_{\text{b}}(k) \\ &= a_{\text{lin}}(\alpha) \cdot x_{\text{b}}(k) + \Delta a(\alpha) \cdot \mathcal{F}^{-1}\{\Delta a(\omega)\} * x_{\text{b}}(k) \\ &\approx a_{\text{lin}}(\alpha) \cdot x_{\text{b}}(k) + P_{\Delta a}(a_{\text{lin}}(\alpha)) \cdot h_{\text{DEQ}}(k) * x_{\text{b}}(k) \end{aligned} \tag{10}$$

where $\mathcal{F}^{-1}(\cdot)$ denotes the inverse Fourier transform. The effect of directional equalizing can be observed in Fig. 3 (b), which displays an almost frequency-invariant behavior over the entire wideband frequency range.

The practical operation of the modified version of the microphone array (Fig. 4) does not significantly differ from the conventional version (Fig. 1): The desired notch angle α is still easily controlled by adapting the scalar factor a_{lin}. As an additional step, the polynomial $P_{\Delta a}(a_{\text{lin}})$ must be evaluated.

5.1 Discussion

The distorted notch curve of the standard differential array (Fig. 3 (a)) not only limits the ability to suppress interfering sound sources, but it can even compromise the accurate NLMS adaptation of the steering angle α (see section 6).

It should be noted that other approaches exist to cope with distorted notch characteristics. For example, a smaller microphone distance D could be used so that the product ωD in Eq. (6) remains sufficiently small. The downside of this approach is a stronger highpass effect of the array which, in turn, requires heavier output equalization with a more pronounced lowpass filter $h_{\mathrm{EQ}}(k)$. In a real system, this leads to a higher amplification of the microphone noise, particularly at low frequencies. For the present example with $D = 1.8\,\mathrm{cm}$, less than half of the original distance is required to obtain a comparably straight directional response. This, however, comes at the cost of a significantly increased noise gain (+10dB) over a wide frequency range.

If a subband (or frequency domain) realization of the differential array is used (see, e.g., H. Puder), the subband steering factors can be individually adapted which, naturally, helps to compensate the notch curve distortion of Fig. 3. The problem has also been identified by M. Buck and M. Rößler, where it is proposed to combine the advantages of the differential array with that of a superdirective endfire array. A more general proposal for frequency-invariant beamforming techniques was made in L. C. Parra, "Steerable frequency-invariant beamforming for arbitrary arrays," Journal of the Acoustical Society of America, Vol. 119, No. 6, pages 3839 to 3847, 2006.

6. Fast Notch Adaptation

The goal of a *notch adaptation algorithm* is to automatically align the notch angle α of the differential array with the incidence angle θ of the (main) interferer. In the common application scenario, the undesired noise or interference sources are assumed to lie in the rear half plane, i.e., $\theta = 90° \ldots 180°$. The standard approach to adapt the factor a (or a_{lin} if directional equalization is used) and therefore the notch angle α is the *(normalized) least mean square* (NLMS) algorithm. The goal here is to minimize the power of the output signal $y(k)$, i.e.

$$E\{(x_f(k) - a \cdot x_b(k))^2\} \rightarrow \min, \qquad (11)$$

where, usually, $0 \leq a \leq 1$, i.e., $180° \geq \alpha \geq 90°$ is enforced. The stepwise update rule of this method is (e.g., H. Puder)

$$a(k) = a(k-1) + \frac{\mu}{\hat{\sigma}^2_{x_\mathrm{b}}} \cdot x_\mathrm{b}(k) \cdot y(k) \tag{12}$$

with the short term power estimate $\hat{\sigma}^2_{x_\mathrm{b}}$ of the backward cardioid signal and the constant stepsize μ.

For our proposed notch adaptation algorithm, which is inspired by the "optimum stepsize NLMS" of M. Pawig, G. Enzner, and P. Vary, "Adaptive sampling rate correction for acoustic echo control in voice-over-IP," IEEE Transactions on Signal Processing, Vol. 58, pages 189 to 199, January 2010, we separate the coherent and the incoherent (ambient) components of the acoustic environment. The array output (before the lowpass equalizer) therefore becomes

$$y(k) = \underbrace{x_\mathrm{f,c}(k) - a \cdot x_\mathrm{b,c}(k)}_{e(k)} + \underbrace{x_\mathrm{f,a}(k) - a \cdot x_\mathrm{b,a}(k)}_{n(k)} \tag{13}$$

This equation represents the error signal of a single-tap adaptive filter with a noisy input. In the context of the differential array, the noise signal $n(k)$ is due to the incoherent (ambient) noise that cannot be suppressed. The coherent contribution to $y(k)$ should ideally be zero. However, due to the instantaneous misadaptation of the factor a, an error signal $e(k)$ appears at the output. According to adaptive filter theory, the optimal (adaptive) stepsize parameter (see P. Vary and R. Martin, "Digital Speech Transmission – Enhancement, Coding and Error Concealment", Chichester: Wiley, 2006, (13.56)) is

$$\mu_\mathrm{opt} = \frac{E\{e^2(k)\}}{E\{n^2(k)\} + E\{e^2(k)\}} \tag{14}$$

From the signals available, the best approximation of $E\{n^2(k)\}$ is the level $\hat{\sigma}^2_y$ of the microphone array's output $y(k)$ while for $E\{e^2(k)\}$, the assumption of a fixed attenuation factor for the backward cardioid signal is made, i.e. $E\{e^2(k)\} \approx \kappa \cdot \hat{\sigma}^2_{x_\mathrm{b}}$. We set $\kappa = 0.01$ (assumed attenuation of 20 dB) and the adaptive stepsize parameter is hence

$$\mu_{\mathrm{opt}} \approx \frac{\kappa\hat{\sigma}_{x_\mathrm{b}}^2}{\hat{\sigma}_y^2 + \kappa\hat{\sigma}_{x_\mathrm{b}}^2} \tag{15}$$

with the recursively estimated short term powers $\hat{\sigma}_{x_\mathrm{b}}^2$ and $\hat{\sigma}_y^2$, which leads to the new NLMS update rule

$$a(k) = a(k-1) + \xi \frac{\kappa}{\hat{\sigma}_y^2 + \kappa\hat{\sigma}_{x_\mathrm{b}}^2} \cdot x_\mathrm{b}(k) \cdot y(k) \tag{16}$$

The adaptation can be deliberately slowed down by the factor $0 \le \xi \le 1$ to avoid artifacts that stem from the single-tap prediction which does not apply any smoothing.

The combination of the proposed NLMS notch adaptation with the directional equalizer of section 5 is straight forward. The equalizer can indirectly influence and enhance the notch adaptation via the array output signal $y(k)$.

Although the frequency dependency of $a(k)$ is disregarded in the NLMS update, this slight error is immediately compensated for at the next sample instant when the DEQ is in turn adapted to the new steering factor $a(k+1)$ (or $a_{\mathrm{lin}}(k+1)$ in this case).

6.1 Evaluation & Discussion

The performance of the proposed fast notch adaptation algorithm is contrasted with the conventional NLMS using a fixed stepsize in Fig. 5. The graph illustrates the adaptation process for a synthetic sound field with a single sound source that arrives from changing angles θ. In the example, only the angle of $\theta = 0°$ is associated with a desired source. In this case, the adaptation should not drift towards the $90°$ boundary but rather maintain the previously identified steering factor a. The underlying assumption is that an interferer does not move while being inactive. The fast version of the constant stepsize NLMS (Eq. (12)) (gray curve) for example drifts towards $90°$ easily in case of activity of the desired sound source, but even the slower version (blue curve) is not able to maintain a once identified steering factor in all situations. The proposed adaptive NLMS (Eq. (16)) (green curve) mostly solves this problem while converging almost instantaneously towards new interfering sound sources.

However, all three approaches exhibit an unstable behavior for $\theta = 135°$. These surprisingly strong fluctuations, which are clearly audible and also disturbing, are explained by intermittent *high frequency sounds* in the interferer source (such as fricatives in speech). If a high frequency sound event occurs from $\theta = 135°$, the adaptation algorithm (which minimizes the output power) shifts the distorted notch characteristic of Fig. 3 (a) to the left, i.e., toward $\alpha = 90°$ or $\alpha_{\text{lin}} = 1$. Low frequency sounds yield the expected result ($\alpha = 135°$ or $\alpha_{\text{lin}} \approx 0.17$) instead. If the DEQ algorithm from section 5 is activated (red curve in Fig. 5), the unstable adaptation behavior vanishes almost completely.

7. Implemented System

Using fixed point arithmetic, the described differential microphone array (including the proposed enhancements) has been implemented on a signal processor of a wireless loudspeaker (*Binauric Boom Boom*) which is, at the same time, a handsfree communication device. Two miniature digital MEMS microphone capsules are placed on the top of the device with $D = 1.8\,\text{cm}$. The microphones offer SNRs of more than 60 dB which open up the possibility of a differential microphone array with a sufficiently low noise level. An example application scenario is a handsfree call in an office where another colleague is working on the opposite side of the desk. The colleague's noise (typing, voice, etc.) can then be canceled out when placing a call with Boom Boom.

The signal processing software has been developed with the help of the RTProc rapid real-time prototyping framework (see H. Krüger and P. Vary) – the developer interface for algorithm parameterization is shown in Fig. 6. Beginning with a Matlab prototype (based on framewise processing), several other versions have been subsequently developed: A parameterizable C version, a C version with generated parameter tables, a C version based on fixed point arithmetic with an emulated instruction set and generated parameter tables, and finally optimized assembler code for the signal processor with generated parameter tables. All versions can be verified against each other and there is the possibility to step back to Matlab and add or modify features. The measured complexity of the assembler code is approximately 7 MIPS for the wideband sampling rate ($f_s = 16\text{kHz}$).

8. Conclusions

A number of algorithmic enhancements for the standard differential microphone according to G. Elko and A.-T. N. Pong have been proposed herein:

- The described input level alignment procedure enables the array to work reliably despite certain assembly and sensor tolerances.
- With the proposed directional equalizer, a frequency invariant notch characteristic can be obtained even for larger microphone distances. A larger distance also helps to confine the noise gain of the array. Therefore, the array becomes practically usable even for higher sampling rates (e.g., $16\mathrm{kHz}$).
- With the fast notch adaptation algorithm, the noise (or interferer) suppression works more reliable in a broader range of acoustic scenarios. Also, the desired source does not compromise the notch adaptation anymore. Moreover, the combination with the directional equalizer leads to a more stable direction of arrival tracking.

The implementation of the proposed techniques in a new commercial handsfree communication device was described. Finally, the subjective listening impression confirms that, with the new algorithms, interfering sound sources are suppressed much more reliably than with the conventional microphone array. The now frequency-invariant notch characteristic is not the only reason for this. Rather, only the combined application of both DEQ and fast notch adaptation facilitates the clearly improved direction of arrival tracking and interference suppression.

Other variations to the disclosed embodiments can be understood and effected by those skilled in the art in practicing the claimed invention, from a study of the drawings, the disclosure, and the appended claims.

In the claims, the word “comprising” does not exclude other elements or steps, and the indefinite article “a” or “an” does not exclude a plurality.

A single unit or device may fulfill the functions of several items recited in the claims. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage.

Any reference signs in the claims should not be construed as limiting the scope.

CLAIMS:

1. A method for generating a directional sound signal ($y(k)$) from first and second sound signals ($x_1(k)$, $x_2(k)$), which are generated by a first and a second microphone (M1, M2), which are separated by a distance (D), wherein the method comprises:

- generating first and second differential sound signals ($x_{\mathrm{f}}(k)$, $x_{\mathrm{b,DEQ}}(k)$) based on the first and second sound signals ($x_1(k)$, $x_2(k)$), and
- generating the directional sound signal ($y(k)$) according to a frequency-dependent directional response pattern based on the first and second differential sound signals ($x_{\mathrm{f}}(k)$, $x_{\mathrm{b,DEQ}}(k)$),

wherein the generating of the second differential sound signal ($x_{\mathrm{b,DEQ}}(k)$) comprises generating a difference signal ($x_{\mathrm{b}}(k)$) of the first and the second sound signals ($x_1(k)$, $x_2(k)$) and a frequency-selective processing that depends on a steering angle (α), which indicates a desired direction of maximum attenuation of the frequency-dependent directional response pattern, wherein the frequency-selective processing adjusts the actual direction of maximum attenuation of the frequency-dependent directional response pattern to correspond to the steering angle (α) substantially independent of frequency (ω) over the frequency range of the directional sound signal ($y(k)$).

2. The method according to claim 1, wherein the frequency-selective processing comprises weighting the difference signal ($x_{\mathrm{b}}(k)$) with an approximated steering factor ($a_{\mathrm{lin}}(\alpha)$) that is independent of frequency (ω) to generate a weighted difference signal ($a_{\mathrm{lin}}(\alpha) \cdot x_{\mathrm{b}}(k)$) and correcting for the approximation by adding a correction signal ($P_{\Delta\alpha}(a_{\mathrm{lin}}(\alpha)) \cdot h_{\mathrm{DEQ}}(k) * x_{\mathrm{b}}(k)$) that is generated from the difference signal ($x_{\mathrm{b}}(k)$) in dependence of frequency (ω) and the steering angle (α).

3. The method according to claim 2, wherein the generation of the correction signal ($P_{\Delta\alpha}(a_{\mathrm{lin}}(\alpha)) \cdot h_{\mathrm{DEQ}}(k) * x_{\mathrm{b}}(k)$) comprises applying two separate operations, one being dependent on frequency (ω) and independent of the steering angle (α) and one being dependent on the steering angle (α) but independent of frequency (ω).

4. The method according to claim 2 or 3, wherein the generation of the correction signal ($P_{\Delta\alpha}(a_{\mathrm{lin}}(\alpha)) \cdot h_{\mathrm{DEQ}}(k) * x_{\mathrm{b}}(k)$) comprises filtering the difference signal ($x_{\mathrm{b}}(k)$) with

a filter ($h_{\mathrm{DEQ}}(k)$) that is dependent on frequency (ω) and independent of the steering angle (α) to generate a filtered difference signal ($h_{\mathrm{DEQ}}(k) * x_{\mathrm{b}}(k)$).

5. The method according to claim 4, wherein the generation of the correction signal ($P_{\Delta a}(a_{\mathrm{lin}}(\alpha)) \cdot h_{\mathrm{DEQ}}(k) * x_{\mathrm{b}}(k)$) further comprises weighting the filtered difference signal ($h_{\mathrm{DEQ}}(k) * x_{\mathrm{b}}(k)$) with a factor ($P_{\Delta a}(a_{\mathrm{lin}}(\alpha)$) that is dependent on the steering angle (α) and independent of frequency (ω).

6. The method according to claim 5, wherein the factor ($P_{\Delta a}(a_{\mathrm{lin}}(\alpha)$) is determined by using a polynomial approximation that is evaluated with the steering angle (α) or the approximated steering factor ($a_{\mathrm{lin}}(\alpha)$).

7. The method according to any of claims 1 to 6, wherein the method further comprises filtering the directional sound signal ($y(k)$) with a low-pass filter ($h_{\mathrm{EQ}}(k)$) to generate a filtered directional sound signal ($y_{\mathrm{EQ}}(k)$).

8. The method according to any of claims 2 to 7, wherein the approximated steering factor ($a_{\mathrm{lin}}(\alpha)$) for a time instance ($k-1$) is adapted for the following time instance (k) by adding an adaptation value that is scaled by a stepsize parameter (μ_{opt}), wherein the stepsize parameter (μ_{opt}) is adapted in dependence of estimated energies of coherent and incoherent sound components.

9. The method according to claim 8, wherein the energy of the incoherent sound components is approximated by the estimated short-term energy ($\hat{\sigma}_y^2$) of the directional sound signal ($y(k)$) and the energy of the coherent sound components is approximated by a fraction ($\kappa\hat{\sigma}_{x_{\mathrm{b}}}^2$) of the estimated short-term energy ($\hat{\sigma}_{x_{\mathrm{b}}}^2$) of the difference signal ($x_{\mathrm{b}}(k)$).

10. The method according to any of claims 1 to 9, wherein the method further comprises estimating a relative gain (g_{c}) of the first and the second microphone (M1, M2) and equalizing power levels of the first and the second microphone (M1, M2) based on the relative gain (g_{c}).

11. The method according to claim 10, wherein the relative gain (g_c) is determined based on recursively estimated variances ($\hat{\sigma}_{x_1}^2$, $\hat{\sigma}_{x_2}^2$) of the first and second sound signals ($x_1(k)$, $x_2(k)$).

12. The method according to any of claims 1 to 11, wherein the first and the second microphone (M1, M2) are omnidirectional microphones.

13. An apparatus (2) for generating a directional sound signal ($y(k)$) from first and second sound signals ($x_1(k)$, $x_2(k)$), which are generated by a first and a second microphone (M1, M2), which are separated by a distance (D), wherein the apparatus (2) comprises:

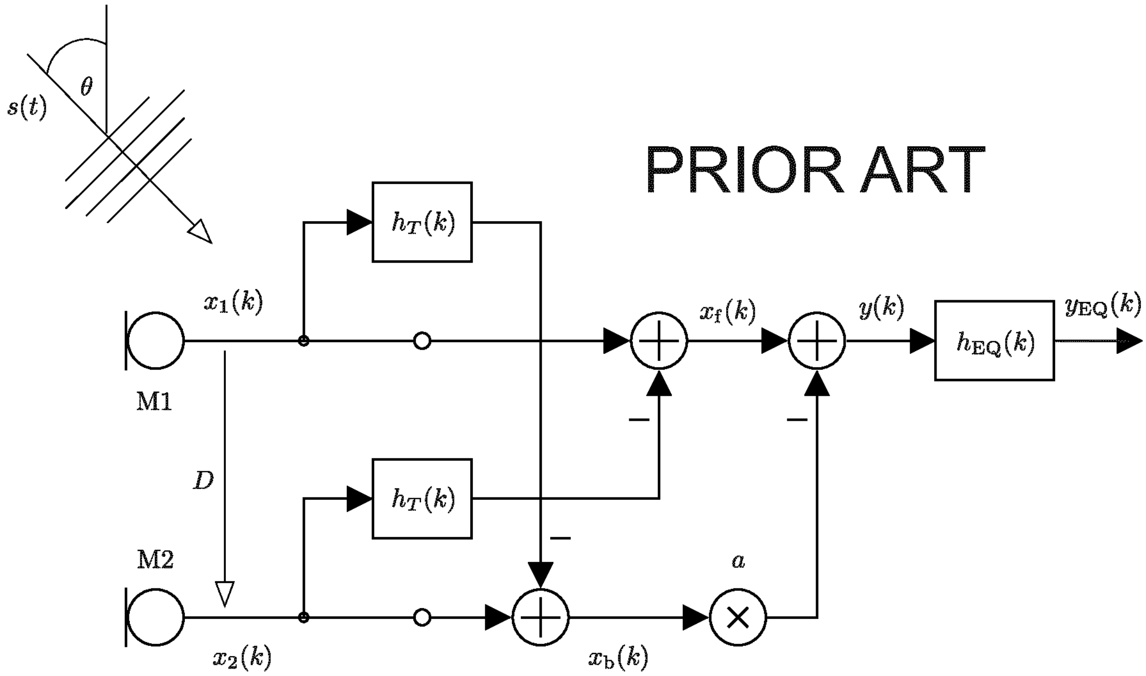
- first generating means for generating first and second differential sound signals ($x_f(k)$, $x_{b,DEQ}(k)$) based on the first and second sound signals ($x_1(k)$, $x_2(k)$), and
- second generating means for generating the directional sound signal ($y(k)$) according to a directional response pattern based on the first and second differential sound signals ($x_f(k)$, $x_{b,DEQ}(k)$),

wherein the generating of the second differential sound signal ($x_{b,DEQ}(k)$) comprises generating a difference signal ($x_b(k)$) of the first and the second sound signals ($x_1(k)$, $x_2(k)$) and a frequency-selective processing that depends on a steering angle (α), which indicates a desired direction of maximum attenuation of the frequency-dependent directional response pattern, wherein the frequency-selective processing adjusts the actual direction of maximum attenuation of the frequency-dependent directional response pattern to correspond to the steering angle (α) substantially independent of frequency (ω) over the frequency range of the directional sound signal ($y(k)$).

14. A system (1), wherein the system (1) comprises:

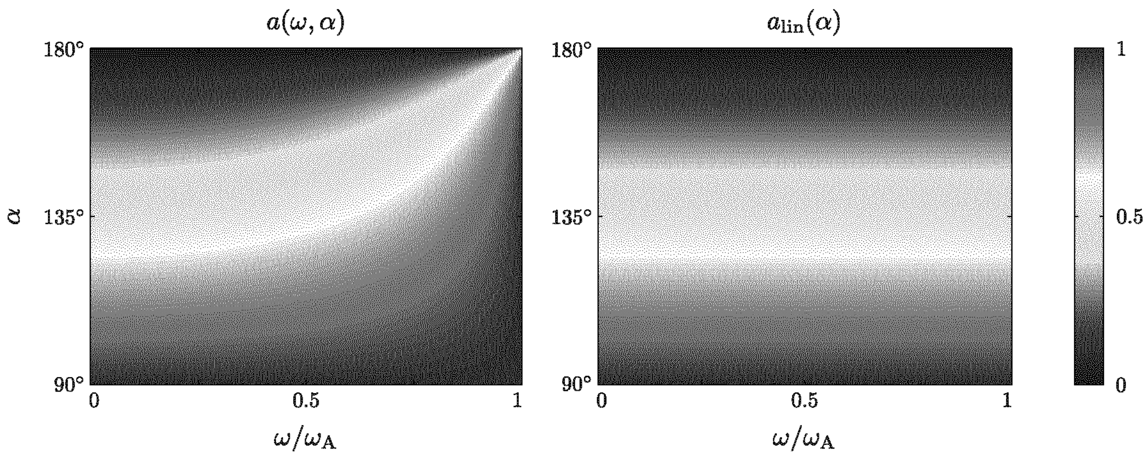
- a first and a second microphone (M1, M2), which are separated by a distance (D) and generate first and second sound signals ($x_1(k)$, $x_2(k)$), and
- an apparatus (2) as defined in claim 13.

15. A computer program comprising program code means, which, when run on a computer controlling the apparatus (2) according to claim 13, perform the steps of the method according to any of claims 1 to 12.



PRIOR ART
$s(t)$
θ
$h_T(k)$
$x_1(k)$
$x_f(k)$
$y(k)$
$h_{EQ}(k)$
$y_{EQ}(k)$
M1
$-$
$-$
D
$h_T(k)$
$-$
M2
a
$x_2(k)$
$x_b(k)$

FIG. 1



$a(\omega, \alpha)$
$a_{lin}(\alpha)$
180°
135°
90°
α
0
0.5
1
ω/ω_A
180°
135°
90°
0
0.5
1
ω/ω_A
1
0.5
0

FIG. 2

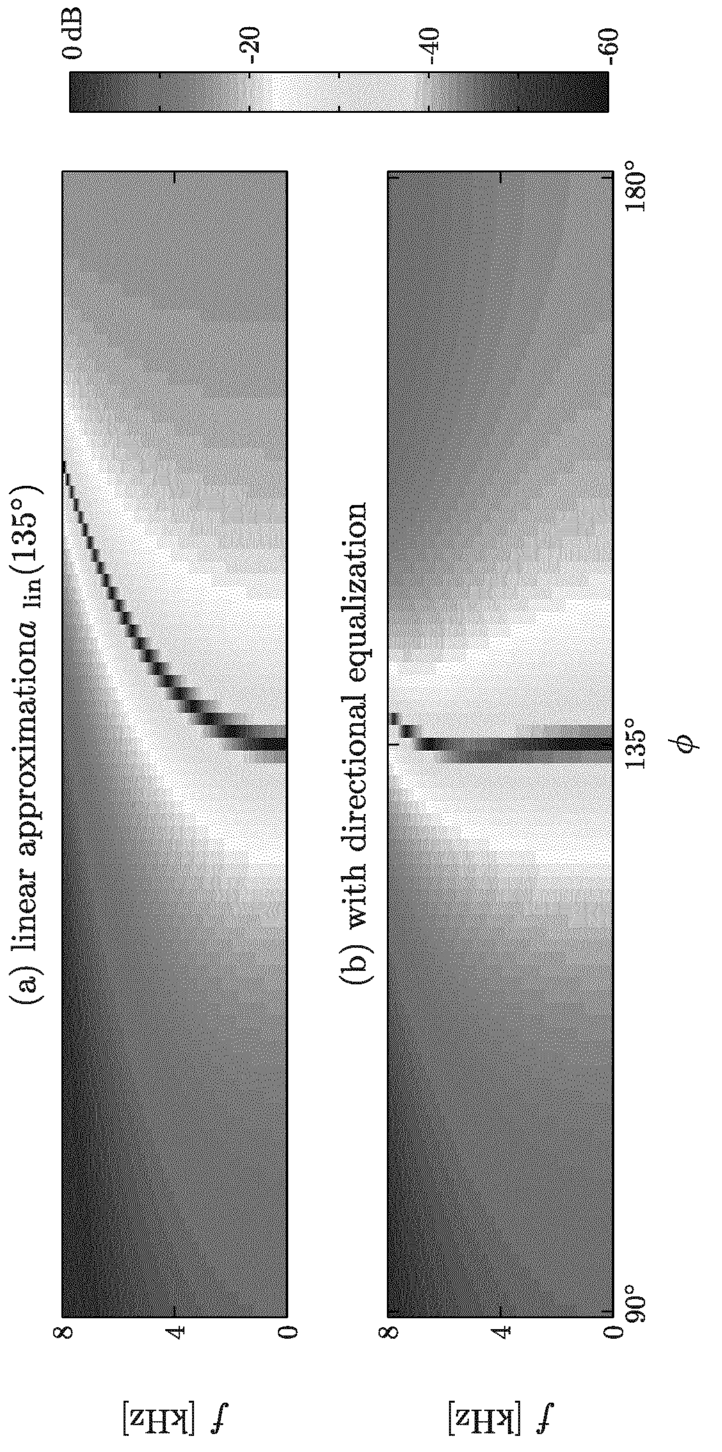
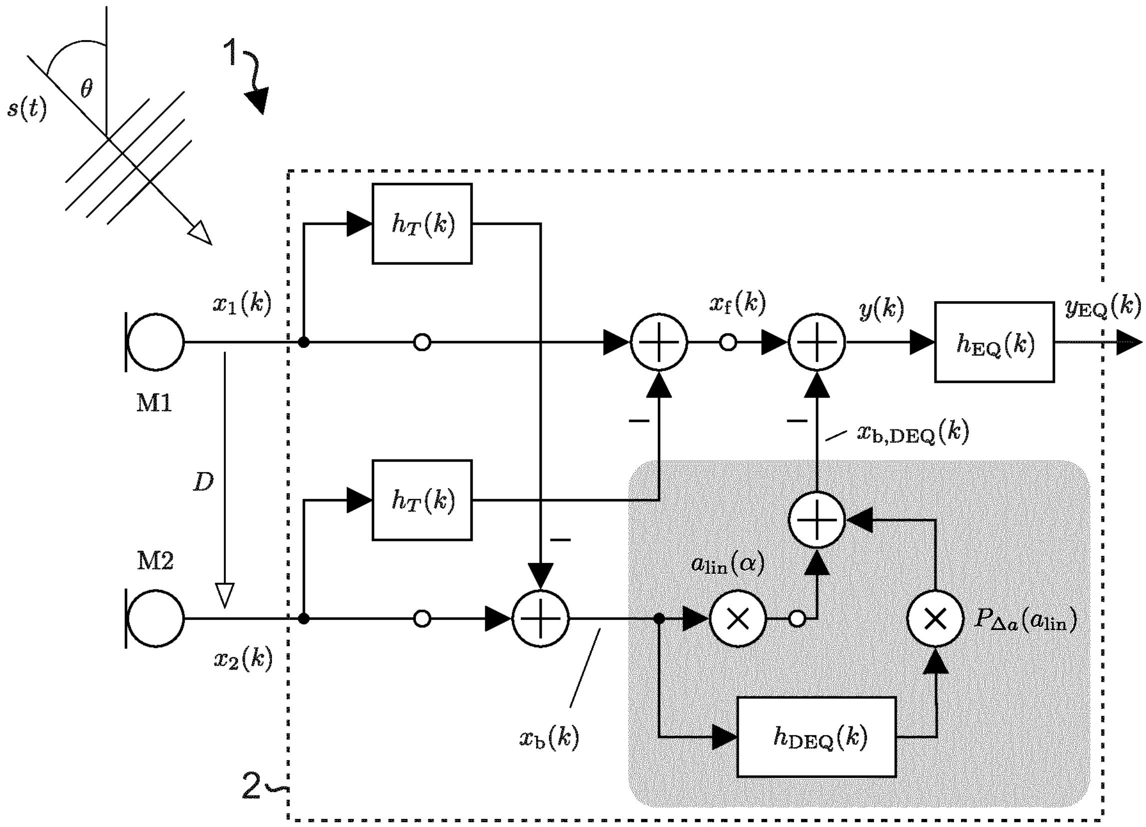

(a) linear approximation $a_{\text{lin}}(135°)$
(b) with directional equalization
f [kHz]
f [kHz]
8
4
0
90°
135°
180°
ϕ
0 dB
-20
-40
-60

FIG. 3



1
2
$s(t)$
θ
M1
M2
D
$x_1(k)$
$x_2(k)$
$h_T(k)$
$h_T(k)$
$x_f(k)$
$y(k)$
$h_{EQ}(k)$
$y_{EQ}(k)$
$x_{b,DEQ}(k)$
$a_{lin}(\alpha)$
$P_{\Delta a}(a_{lin})$
$x_b(k)$
$h_{DEQ}(k)$

FIG. 4

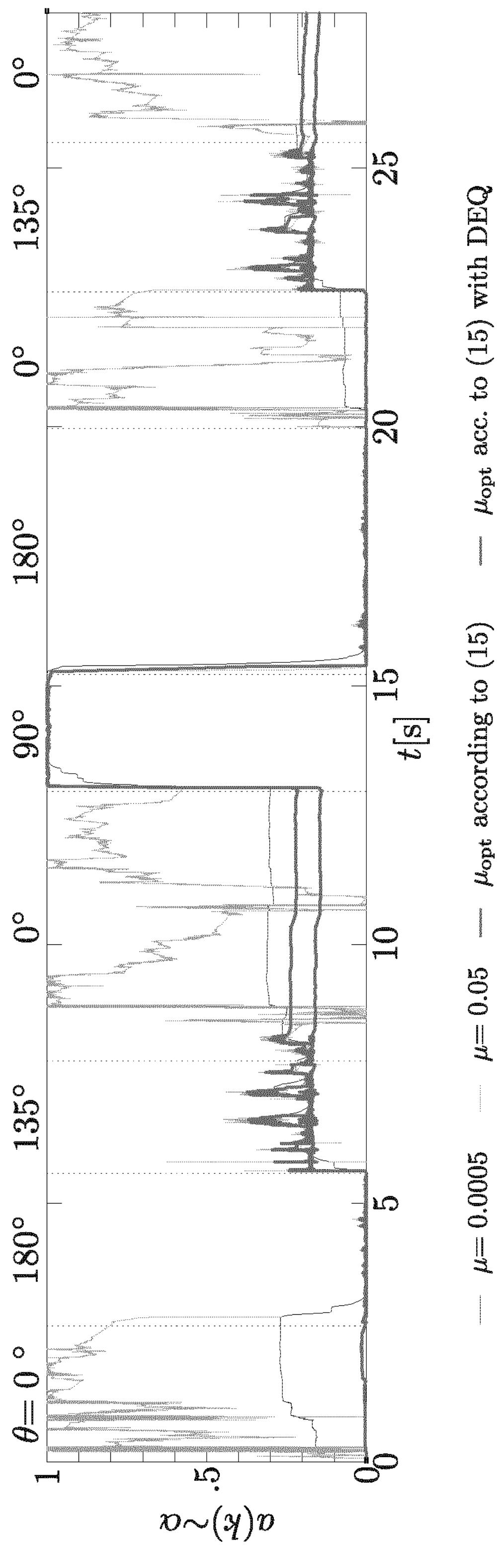

$\theta = 0°$
180°
135°
0°
90°
180°
0°
135°
0°
0
5
10
15
20
25
t[s]
1
.5
0
$\alpha(k) \sim \alpha$
$\mu = 0.0005$
$\mu = 0.05$
μ_{opt} according to (15)
μ_{opt} acc. to (15) with DEQ

FIG. 5

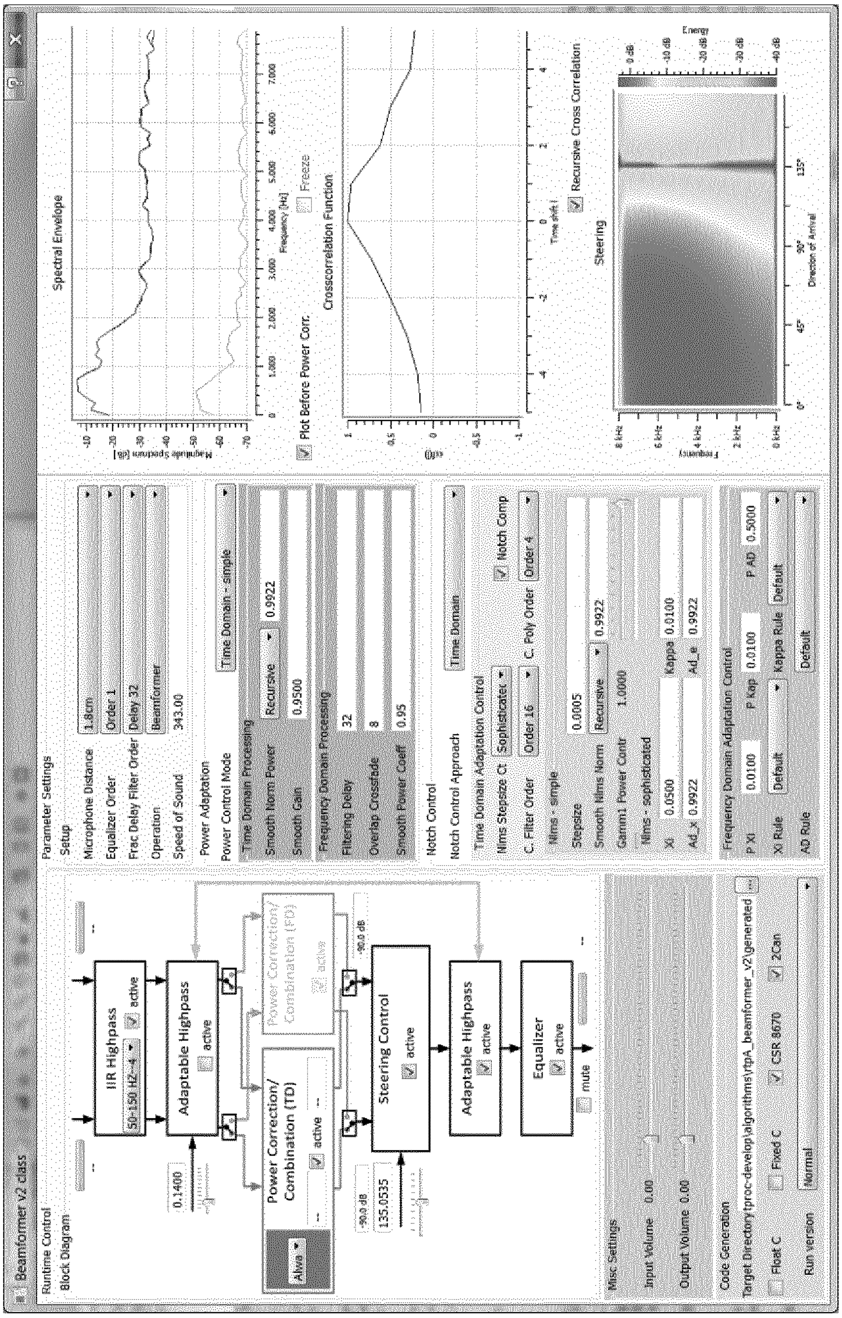

Beamformer v2 class
Runtime Control
Block Diagram
IIR Highpass
Adaptable Highpass
Power Correction/ Combination (TD)
Power Correction/ Combination (FD)
Steering Control
Equalizer
Misc Settings
Input Volume
Output Volume
Code Generation
Parameter Settings
Setup
Microphone Distance
Equalizer Order
Frac Delay Filter Order
Operation
Speed of Sound
Power Adaptation
Power Control Mode
Time Domain Processing
Frequency Domain Processing
Notch Control
Notch Control Approach
Time Domain Adaptation Control
Frequency Domain Adaptation Control
Spectral Envelope
Crosscorrelation Function
Steering
Recursive Cross Correlation
Direction of Arrival

FIG. 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/070243

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04R3/00
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04R

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

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

EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ELKO G W ET AL: "A simple adaptive first-order differential microphone", APPLICATIONS OF SIGNAL PROCESSING TO AUDIO AND ACOUSTICS, 1995., IEEE ASSP WORKSHOP ON NEW PALTZ, NY, USA 15-18 OCT. 1995, NEW YORK, NY, USA,IEEE, US, 15 October 1995 (1995-10-15), pages 169-172, XP010154658, DOI: 10.1109/ASPAA.1995.482983 ISBN: 978-0-7803-3064-1 the whole document -----	1-15

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 June 2015	16/06/2015

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Coda, Ruggero

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	BUCK M: "ASPECTS OF FIRST-ORDER DIFFERENTIAL MICROPHONE ARRAYS IN THE PRESENCE OF SENSOR IMPERFECTIONS", EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS, WILEY & SONS, CHICHESTER, GB, vol. 13, no. 2, 1 March 2002 (2002-03-01), pages 115-122, XP001123749, ISSN: 1124-318X page 115 - page 119 -----	1-15
A	ELKO G W ET AL: "A steerable and variable first-order differential microphone array", IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 1997. ICASSP-97, MUNICH, GERMANY 21-24 APRIL 1997, LOS ALAMITOS, CA, USA,IEEE COMPUT. SOC; US, US, vol. 1, 21 April 1997 (1997-04-21), pages 223-226, XP010226175, DOI: 10.1109/ICASSP.1997.599609 ISBN: 978-0-8186-7919-3 the whole document -----	1-15
A	DERKX R M M ET AL: "Theoretical Analysis of a First-Order Azimuth-Steerable Superdirective Microphone Array", IEEE TRANSACTIONS ON AUDIO, SPEECH AND LANGUAGE PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, USA, vol. 17, no. 1, 1 January 2009 (2009-01-01), pages 150-162, XP011241214, ISSN: 1558-7916, DOI: 10.1109/TASL.2008.2006583 the whole document -----	1-15
A	WO 2014/062152 A1 (MH ACOUSTICS LLC [US]) 24 April 2014 (2014-04-24) Audio signal processing method for cell phones, involves determining portion of signals resulting from incoherence and sources having propagation speeds different from acoustic signals, and filtering signals; page 3, line 20 - page 30, line 20 -----	1-15

-/--

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2014/070243

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	OSAMU HOSHUYAMA ET AL: "A Robust Adaptive Beamformer for Microphone Arrays with a Blocking Matrix Using Constrained Adaptive Filters", IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE SERVICE CENTER, NEW YORK, NY, US, vol. 47, no. 10, 1 October 1999 (1999-10-01), XP011058725, ISSN: 1053-587X the whole document -----	1-15

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2014/070243

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2014062152 A1	24-04-2014	EP 2848007 A1	18-03-2015
		WO 2014062152 A1	24-04-2014