| (12) | PATENT | (11) Application No. **AU 199883642 B2** |
| (19) | **AUSTRALIAN PATENT OFFICE** | (10) Patent No. **747618** |

(51) International Patent Classification ^{6} : G10K 11/178, H04M 9/08	A1	(11) International Publication Number: **WO 99/03091**
		(43) International Publication Date: 21 January 1999 (21.01.99)

(54) Title: METHODS AND APPARATUS FOR MEASURING SIGNAL LEVEL AND DELAY AT MULTIPLE SENSORS
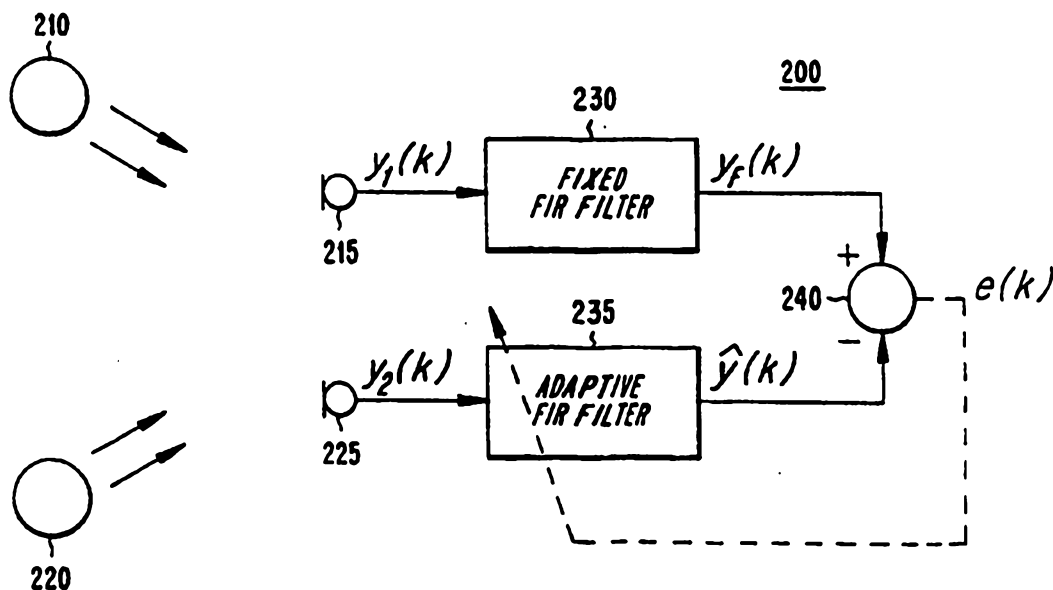
(57) Abstract

Methods and apparatus for automatic gain and time delay computations between multiple sensors. In exemplary embodiments, one sensor is treated as a reference sensor with respect to a measured signal quantity which is assumed to be a narrowband process. The relative gain and relative time delay in the additional sensors with respect to the same measured quantity are then automatically calculated based on an adaptive filtering algorithm. Advantageaously, the disclosed embodiments are implemented using standard digital signal processing components.

-1-

METHODS AND APPARATUS FOR MEASURING
SIGNAL LEVEL AND DELAY AT MULTIPLE SENSORS

BACKGROUND

The present invention relates to signal processing, and more particularly to

the measurement of signal levels and time delays at multiple signal sensors.

In many signal processing applications, it is desirable to determine the

relative sensitivity of multiple signal sensors with respect to a particular signal

source. For example, in the context of hands-free mobile telephony, dual

microphones can be used in combination with beamforming methods to reduce the

effects of background noise and echoes in an automobile. To do so, information

regarding the relative sensitivities of the microphones with respect to different

acoustic sources is used, for example, to form a spatial beam toward a particular

user and/or to form a spatial notch against another user or a loudspeaker. Such an

approach requires that dynamic information with respect to microphone sensitivity

be quickly and accurately obtained.

Figure 1 depicts a prior art system 100 for measuring the relative

sensitivities of dual microphones with respect to different signal sources in the

context of hands-free mobile telephony. As shown, the prior art system 100

includes a first microphone 115, a second microphone 125, an adaptive filter 135

and a summing device 140. An output $y_1(k)$ of the first microphone 115 is coupled

to a positive input of the summing device 140, and an output $y_2(k)$ of the second

microphone 125 is coupled to an input of the adaptive filter 135. An output $\hat{y}_1(k)$ of

the adaptive filter 135 is coupled to a negative input of the summing device 140,

and an output $e(k)$ of the summing device 140 is used as a feedback signal to the

adaptive filter 135.

As shown, the first microphone 115 is positioned nearer a first source 110,

and the second microphone 125 is positioned nearer a second source 120. For

-2-

example, the first microphone 115 can be a hands-free microphone attached to a

sun visor situated nearer a driver of an automobile, and the second microphone

125 can be a built-in microphone within a mobile unit attached nearer a passenger

in the automobile. Although it is not shown in Figure 1, those skilled in the art

5   will appreciate that analog pre-processing and analog-to-digital conversion

circuitry can be included at the output of each of the first and second microphones

115, 125 so that digital signals are processed by the adaptive filter 135 and the

summing device 140. The output $e(k)$ of the summing device 140 represents the

difference between the output $y_1(k)$ of the first microphone 115 and the output $\hat{y}_1(k)$

10   of the adaptive filter 135 and is referred to herein as an error signal.

In operation, filter coefficients of the adaptive filter 135 are adjusted using

a least-squares algorithm such that the error signal $e(k)$ is minimized. In other

words, the adaptive filter 135 is adjusted such that the output $\hat{y}_1(k)$ of the adaptive

filter 135 is as close as possible to (i.e., is an estimator of) the output $y_1(k)$ of the

15   first microphone 115. Thus, the adaptive filter 135 attempts to model the signal

effects created by the physical separation of the microphones 115, 125. For

example, when the passenger 120 is speaking, his or her voice will reach the first

microphone 115 slightly later than it will reach the second microphone 125, and

the corresponding speech signal level received at the first microphone 115 will be

20   somewhat attenuated as compared to the level received at the second microphone

125. Thus, the adaptive filter 135 is adjusted to provide similar delay and

attenuation effects.

As a result, the relative time delay and signal attenuation at the

microphones with respect to each user can be calculated based on the coefficients

25   of the adaptive filter 135 as described, for example, in Y.T. Chan, J.M. Riley and

J.B. Plant, "A parameter estimation approach to time delay estimation and signal

detection", IEEE Transactions on Acoustics, Speech and Signal Processing, vol.

ASSP-28, Feb. 1980, which is incorporated herein in its entirety by reference.

One disadvantage of the system of Figure 1, however, is that its performance deteriorates significantly in the presence of background noise. As a result, the system of Figure 1 is not useful in most practical applications, where significant background noise (e.g., road and traffic noise) is commonplace. Thus, there is a

5   need for improved methods and apparatus for measuring relative signal levels and time delays at multiple sensors.

SUMMARY OF THE INVENTION

According to a first aspect of the invention there is provided a signal processing device, including:

10   a first signal sensor;

a first filter having an input coupled to an output of said first sensor;

a second signal sensor;

a second filter having an input coupled to an output of said second sensor and having an adjustable filtering characteristic;

15   a summing device having a first input coupled to an output of said first filter and a second input coupled to an output of said second filter, wherein said summing device outputs a difference between the first and second inputs, and wherein the adjustable filtering characteristic of said second filter is adjusted in dependence upon the output of said summing device; and

20   a processor for computing an estimate of at least one parameter describing a relationship between said first and second sensors, wherein the at least one parameter indicates at least one of a relative time delay and a relative scale factor between said first and second sensors with respect to a signal source, and wherein the estimate is computed as a function of a filtering

25   characteristic of said first filter and as a function of the adjustable filtering characteristic of said second filter.

According to a second aspect of the invention there is provided a method for processing signals, including the steps of:

detecting a first signal using a first signal sensor;

30   filtering the first signal using a first filter to provide a first filtered signal;

detecting a second signal using a second signal sensor;
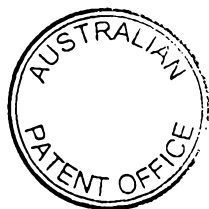
filtering the second signal using a second filter to provide a second filtered signal;

computing a difference between the first filtered signal and the second filtered signal;

5   adjusting a filtering characteristic of the second filter in dependence upon the difference obtained in said step of computing; and

estimating at least one parameter as a function of a filtering characteristic of said first filter and as a function of the filtering characteristic of said second filter, wherein the at least one parameter indicates at least one of a relative time

10   delay and a relative scale factor between said first and second sensors with respect to a signal source.

The present invention addresses the above-described and other needs by providing a system in which a fixed filter and an adaptive filter are used in combination to provide accurate and robust estimates of signal levels and time

15   delays for multiple sensors. In exemplary embodiments, the fixed filter includes at least one relatively narrow passband which is used to distinguish signal sources of interest from broad-band background noise. In the embodiments, the fixed filter is coupled to a reference sensor and the adaptive filter is coupled to a secondary sensor. An error signal derived from the outputs of the fixed filter and

20   the adaptive filter is used to adjust filter coefficients of the adaptive filter according to a suitable least-squares algorithm. The coefficients of the fixed filter and the adaptive filter are used to compute estimates of the time delay and relative level between the two sensors. The estimates can then be used to make decisions regarding sensor selection and beamforming.

25   In exemplary embodiments, the functionality of the system is supplemented with an activity detector which indicates when no signal of interest is present. In the activity detector, accumulated energy in the adaptive filter is compared with an expected least value derived from the coefficients of the fixed filter. When the

30

-4-

accumulated energy is smaller than the expected value, indicating that there is no signal of interest present (i.e., only background noise is present), the time delay and relative level estimates are set to appropriate values to ensure proper operation of the system even during periods where no signals of interest are present.

5   In additional embodiments, more than two signal sensors are employed. In such embodiments, one sensor is treated as a reference sensor and coupled to a fixed filter, while each of the additional sensors is coupled to an adaptive filter. For each additional sensor, an error signal derived from the outputs of the fixed filter and the corresponding adaptive filter is used to update the coefficients of the

10   corresponding adaptive filter. Thus, robust estimates of the time delay and relative signal level between the reference sensor and each additional sensor can be computed, and sophisticated decisions can be made with respect sensor selection and beamforming.

           Generally, the present invention provides a computationally simple yet

15   accurate and robust method for estimating the time delays and relative signal levels at multiple sensors. The teachings of the invention are applicable in a wide variety of signal processing contexts. For example, in addition to the hands-free mobile telephony application described above, the invention may be used for other acoustic applications such as teleconferencing. Additionally, the present invention

20   is applicable in radio communication applications where the signals of interest are radio-frequency transmissions (e.g., from mobile units and/or base stations in a cellular radio system) and the sensors are radio-frequency-sensitive antenna elements. These and other features and benefits of the present invention are explained hereinafter with reference to the illustrative examples shown in the

25   accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 depicts the prior art signal level and delay measurement system described above.

Figure 2 depicts a signal level and delay measurement system constructed in accordance with the present invention.

Figure 3 depicts relative signal levels and time delays of two signals detected at dual signal sensors.

Figure 4 depicts an alternate signal level and delay measurement system constructed in accordance with the present invention.

Figure 5 depicts magnitude and phase responses of an exemplary signal filter which can be employed in the exemplary systems of Figures 2 and 4.

Figure 6 depicts exemplary speech and noise signals which are used to demonstrate operation of exemplary embodiments of the present invention.

Figure 7 depicts signal level and delay estimates generated by an exemplary embodiment of the present invention based on the signals of Figure 6.

DETAILED DESCRIPTION OF THE INVENTION

Figure 2 depicts a level and delay measurement system 200 constructed in accordance with the teachings of the present invention. As shown, the system 200 includes a first sensor 215, a second sensor 225, a fixed FIR filter 230, an adaptive FIR filter 235 and a summing device 240. An output $y_1(k)$ of the first sensor 215 is coupled to an input of the fixed filter 230, and an output $y_F(k)$ of the fixed filter 230 is coupled to a positive input of the summing device 240. An output $y_2(k)$ of the second sensor 225 is coupled to an input of the adaptive filter 235 and an output $\hat{y}(k)$ of the adaptive filter 235 is coupled to a negative input of the summing device 240. An error signal $e(k)$ which is output by the summing device 240 is fed back to the adaptive filter 235.

-6-

As shown, the first sensor 215 is positioned nearer a first signal source 210, and the second sensor 225 is positioned nearer a second signal source 220. For example, the first sensor 215 can be a hands-free microphone attached to a sun visor situated nearer a driver of an automobile, and the second sensor 225 can

5   be a built-in microphone within a mobile unit attached nearer a passenger in the automobile. Alternatively, the first and second sensors 215, 225 can be antenna elements positioned nearer first and second radio-frequency signal sources, respectively. Although it is not shown in Figure 2, those skilled in the art will appreciate that analog pre-processing and analog-to-digital conversion circuitry

10   can be included at the output of each of the first and second sensors 215, 225 so that digital signals are processed by the fixed filter 230, the adaptive filter 235 and the summing device 240.

The fixed filter 230 is designed to include at least one relatively narrow pass-band of interest. For example, in the mobile telephony context, a pass-band

15   can correspond to the 300-600 Hz frequency band in which most of the energy of human speech is concentrated. In a radio communication application, a pass-band can correspond to a bandwidth allocated for radio-frequency transmissions. In any case, the coefficients of the fixed filter 230 can be adjusted as necessary to compensate for changes in application requirements or environmental conditions.

20   For example, in a hands-free mobile telephone application, the fixed filter 230 can be set to optimize received signal-to-noise ratio for a particular automobile installation. Furthermore, the coefficients of the filter 230 can be adjusted dynamically, for example in dependence upon measured signal-to-noise ratio.

According to the invention, the fixed filter 230 is designed to provide unity

25   gain and zero phase in each passband. Additionally, the noise gain of the fixed

filter 230 is minimized in order to ensure maximal stop-band attenuation. As described in more detail below, the prior information provided by the fixed filter 230 (i.e., the narrowband nature of the signals output by the fixed filter 230) is used to make the system robust against background noise.

5   In operation, filter coefficients of the adaptive filter 235 are adjusted using a suitable least-squares algorithm such that the error signal $e(k)$ is minimized and such that the output $\hat{y}(k)$ of the adaptive filter 235 is as close as possible to the output $y_F(k)$ of the fixed filter 230. As described below, the relative time delay and signal attenuation at the first and second sensors 215, 225 with respect to each

10   source 210, 220 are calculated based on the coefficients of the adaptive filter 235 and the prior information associated with the fixed filter 230. Although not explicitly shown in Figure 2, those skilled in the art will appreciate that an appropriate digital signal processor can be integrated with the system 200 to perform the least-squares update of the adaptive filter 235 and to compute the time

15   delay and signal level estimates.

In order to clarify the operation of the system 200 of Figure 2, a rigorous mathematical analysis is developed below with respect to Figures 3 and 4. Although the analysis is explicitly developed for two sensors and two signal sources, those skilled in the art will appreciate that the described approach is

20   readily applicable to applications including an arbitrary number of signal sources and sensors. Additionally, although reference is sometimes made to the acoustic hands-free mobile telephony application described above, those skilled in the art will appreciate that the described approach is also applicable to many other signal processing contexts including the radio communications applications previously

25   mentioned.

Figure 3 depicts a typical example of source and sensor placement in two dimensions. In the figure, first and second sensors 215, 225 are positioned adjacent two signal sources 210, 220. As shown, a signal emanating from the first

-8-

signal source 210 (as indicated by a first dashed arc 315) will impinge upon the
first signal sensor 215 before impinging upon the second signal sensor 225. Thus,
the signal received at the second sensor 225 due to the first signal source 210 will
be a delayed and attenuated version of the signal received at the first sensor 215

5   due to the same source 210. Additionally, a signal emanating from the second
source 220 (as indicated by a second dashed arc 325) will impinge upon the
second sensor 225 before impinging upon the first sensor 215, and the signal
received at the first sensor 215 due to the second signal source 220 will be a
delayed and attenuated version of the signal received at the second sensor 225 due

10   to the same source 220. The spacial separation (and thus the corresponding time
delay and level attenuation) of the sensors 215, 225 with respect to the first and
second signal sources 210, 220 are indicated in Figure 3 by second and first line
segments 320, 310, respectively.

       If the first and second sensor inputs (after analog pre-processing and

15   analog-to-digital conversion) at time instant k are denoted by $x_1(k)$ and $x_2(k)$,
respectively, then the second sensor input $x_2(k)$ is generally a delayed and scaled
version of the first sensor input $x_1(k)$. In other words, $x_2(k) = 1/S \cdot x_1(k-\bar{D})$, where
the scale factor S is greater than zero and where the delay $\bar{D}$ may take positive as
well as negative values. Strictly speaking, for $\bar{D} < 0$ (e.g., for signals emanating

20   from the second signal source 220), the first input $x_1(k)$ is a delayed and scaled
version of the second input $x_2(k)$. However, in order to simplify notation, the
second input $x_2(k)$ is denoted the delayed signal for all values of $\bar{D}$ without loss of
generality.

       In order to provide a causal filtering problem, a fixed delay Δ can be

25   introduced into the signal path following the first sensor. While this is a natural

-9-

approach in most applications, it is not a prerequisite for the invention to work as intended. This fact is explained in greater detail below. With the extra delay Δ introduced, one can define first and second intermediate signals $y_1(k)$, $y_2(k)$ as follows:

$$y_1(k) = x_1(k-\Delta) \qquad (1)$$

$$y_2(k) = x_2(k) = \frac{1}{S} q^{\Delta-\bar{D}} y_1(k) = \frac{1}{S} y_1(k+D) \qquad (2)$$

5   where q denotes the well known delay operator (i.e., $qy(k)=y(k+1)$, $q^{-1}y(k)=y(k-1)$, etc.), and where D is defined to be $\Delta-\bar{D}$. Note that, for causal filtering, $\Delta > \bar{D}$.

To aid discussion, Figure 4 illustrates the input signals $x_1(k)$, $x_2(k)$ and the intermediate signals $y_1(k)$, $y_2(k)$ in the context of a level and delay measurement

10   system. The system 400 of Figure 4 is identical to the system 200 of Figure 2 except that a delay block 410 (corresponding to the fixed delay Δ described above) is positioned between the first sensor 215 and the fixed filter 230. In the discussion that follows, it is assumed that the coefficients of the fixed filter 230 are stored in a first coefficient vector c_0 and that the time-varying coefficients of

15   the adaptive filter 235 are stored in a second coefficient vector $\hat{c}(k)$.

Generally, the present invention provides a computationally simple yet accurate method for estimating the delay D and the scale factor S based on the measured sensor inputs $x_1(k)$ and $x_2(k)$. Advantageously, the method is robust against background noise so that it may be used successfully, for example, in the

20   above described hands-free mobile telephony context. The estimated quantities,

say D_k and $\hat{S}_k$ (where k indicates that sensor inputs up to and including time instant k are used for the calculation of D and S), can be used to improve system performance.

For example, in the context of mobile telephony, the estimates D_k and $\hat{S}_k$ can be used in combination with well known beamforming techniques to electronically enhance and reduce the sensitivity of the sensors 215, 225 with respect to the first and second sources 210, 220. For instance, when a particular source is active (e.g., when the driver is speaking), a beam may be formed in the direction of that source to optimize its reception. Alternatively, when a source provides a signal that should be blocked from the sensors (e.g., when a source is a loudspeaker that may cause bothersome feedback or echo), spatial filtering can be employed to diminish the sensitivity of the sensors with respect to that source.

Furthermore, the system can selectively transmit only the signal detected at a particular sensor when a particular source is active. For example, if one sensor is much more sensitive to the passenger than to the driver (e.g., due to a close physical proximity to the passenger), then it may be desirable to transmit only the signal received at that sensor when only the passenger is speaking.

Returning to Figure 4, the signal $y_F(k)$ output by the fixed filter 230 (i.e., the filtered version of the first intermediate signal $y_1(k)$) is given by:

$$y_F(k) = \mathbf{Y}_1(k)^T \mathbf{c}_0 \tag{3}$$

$$\mathbf{Y}_1(k) = (y_1(k) \cdots y_1(k-L))^T \tag{4}$$

$$\mathbf{c}_0 = (c_0 \cdots c_L)^T \tag{5}$$

-11-

where L is the order of the fixed filter 230, and where $\{c_i\}$, $\ell = 0,...,$ L are the

fixed filter coefficients. Additionally, the signal $\hat{y}(k)$ output by the adaptive filter

235 (i.e., the filtered version of the second intermediate signal $y_2(k)$) is given by:

$$\hat{y}(k) = \mathbf{y}_2(k)^T \hat{c}(k-1) \tag{6}$$

$$\mathbf{y}_2(k) = (y_2(k) \cdots y_2(k-L))^T \tag{7}$$

$$\hat{c}(k-1) = (\hat{c}_1(k-1) \cdots \hat{c}_L(k-1))^T \tag{8}$$

where the vector $\hat{c}(k)$ contains the time varying filter coefficients of the adaptive

filter 235. The vector $\hat{c}(k)$ is updated based on the error signal $e(k)$ as follows:

$$e(k) = y_F(k) - \hat{y}(k) \tag{9}$$

$$\hat{c}(k) = \hat{c}(k-1) + \mu \frac{\mathbf{y}_2(k)}{\|\mathbf{y}_2(k)\|^2} e(k) \tag{10}$$

where μ is a gain factor (constant or time-varying) in the interval $0 \le \mu < 2$, and

where $\|\cdot\|^2$ denotes the squared Euclidian vector norm. The adaptive algorithm

described by equations (9) and (10) is the well known Normalized Least Mean

Squares (N-LMS) algorithm. Alternative adaptive schemes, such as the Recursive

Least Squares (RLS) algorithm or the Least Mean Squares (LMS) algorithm can

also be used. For a more detailed description of adaption algorithms generally,

see for example B. Widrow and S.D. Stearns, *Adaptive Signal Processing*,

Prentice Hall, Englewood Cliffs, NJ, 1985, and L. Ljung and T. Söderström,

-12-

Theory and Practice of Recursive Identification, M.I.T. Press, Cambridge, MA, 1983, each of which is incorporated herein by reference. Advantageously, each of the above defined quantities can be computed using standard digital signal processing components.

5   For a broadband source impinging on the sensors 215, 225, the coefficients of the adaptive filter 235 converge toward a delayed and scaled version of the coefficients of the fixed filter 230. In particular, if every coefficient of the fixed vector c_0 is 1 (i.e., if effectively no fixed filter is used), the time-varying vector $\hat{c}(k)$ converges toward an approximation of the scaled delay (i.e., $Sq^{D-\Delta} = Sq^{-D}$).

10   Such a result has been used to estimate time-delays in prior art systems. See, for example, Y.T. Chan, J.M.F. Riley and J.B. Plant, "Modeling of time delay and its application to estimation of nonstationary delays", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-29, No. 3, pp. 577-581, June 1981, which is incorporated herein by reference. One disadvantage

15   associated with such systems is that overall system performance deteriorates significantly when background noise is present, making the systems impractical in most real-world applications.

           Advantageously, the present invention teaches that by incorporating prior knowledge which distinguishes the source signals from background noise, system

20   performance can be significantly improved. To ensure improved overall performance, the priors should be true in all situations. For example, the present invention teaches that such prior information is available when the energy in the source signals of interest is concentrated around one or more center frequencies, while the background noise has a relatively flat and broadband frequency content,

25   or power spectral density. In such a context, the present invention teaches that the fixed FIR filter 230 can be designed as a band-pass filter having one or several pass bands.

-13-

For example, for speech signals in a mobile hands-free scenario, it is reasonable to assume that the energy of the speech signals is concentrated in the interval 100-250Hz. More specifically, the fundamental frequency of a male speaker is typically around 100Hz, and the fundamental frequency of a female

5   speaker is typically around 250Hz. In view of this information, the present invention teaches several possible designs alternatives for the fixed filter 230. For example, the fixed filter 230 can be designed to include two pass-bands, the first and second passbands having center frequencies of 100Hz and 250Hz, respectively. Alternatively, the fixed filter 230 can be designed to include a single

10   pass-band having a center frequency of 200Hz and spanning a frequency band which includes the fundamental frequency of female speakers as well as the first harmonic frequency of male speakers.

In practice, the former approach requires the use of higher order filters as compared to the latter approach. Generally, if the number of design frequencies

15   are doubled, then the order L of the filter is doubled as well. In the discussion that follows, a fixed filter 230 having m distinct pass-bands is considered. At the center frequencies $\{\omega_\ell\}$, $\ell = 1,...,m$, the filter is designed to provide unity gain and zero phase. Additionally, the fixed filter 230 is designed to provide maximal attenuation in the stop-bands by minimizing the filter noise gain (NG) which is

20   defined as:

$$NG = \frac{1}{2\pi i} \oint C_0(z) C_0(z^{-1}) \frac{dz}{z} \qquad (11)$$

-14-

where $C_0(z^{-1}) = c_0 + c_1 z^{-1} + \cdots + c_L z^{-L}$, and where the integration proceeds around the unit circle. By Parseval relation, the noise gain for FIR filters is given by:

$$NG = \sum_{\ell=0}^{L} c_\ell^2. \tag{12}$$

To design the fixed filter 230, consider an input signal $y_{in}(k)$ comprising a sum of sinusoids as follows:

$$y_{in}(k) = \sum_{\ell=1}^{m} \alpha_\ell \sin(\omega_\ell k + \phi_\ell) \tag{13}$$

5 where $\{\omega_\ell\}, \ell = 1, \ldots, m$ are the desired center frequencies of the fixed FIR filter, $\omega_\ell \in (0, \pi)$, $\omega_\ell \neq \omega_j, \ell \neq j$, $\{\alpha_\ell\}$ are unknown constants $\alpha_\ell > 0, \ell = 1, \ldots, m$, and $\{\phi_\ell\}$ are uniformly distributed random variables $\phi_\ell \in (-\pi, \pi], \ell = 1, \ldots, m$. Next consider a fixed FIR filter 230 having a coefficient vector $\mathbf{c}_d$ and providing an output $y_{out}(k)$ which is an exact d-step prediction (after any initial transients have

10 decayed) of the input $y_{in}(k)$ for any d, $-\infty < d < \infty$, as follows:

$$y_{out}(k+d) = \mathbf{y}_{in}(k)^T \mathbf{c}_d \tag{14}$$

$$\mathbf{y}_{in}(k) = (y_{in}(k) \cdots y_{in}(k-L))^T \tag{15}$$

$$\mathbf{c}_d = (c_0 \cdots c_L)^T \tag{16}$$

For $d = 0$ (i.e., for a coefficient vector $\mathbf{c}_0$), the fixed FIR filter 230 provides unity gain and zero phase at the center frequencies $\{\omega_\ell\}, \ell = 1, \ldots, m$ as desired. Additionally, if the sensitivity to broadband noise is minimized (i.e., if the quantity in equation (12) is minimized), then (for a filter length L, such that

15 $L > 2m - 1$) the following result holds true:

-15-

$$\mathbf{c}_d = \mathbf{L}^T(\mathbf{LL}^T)^{-1}\mathbf{p}(d) \tag{17}$$

where **L** is the $2m \times (L + 1)$ matrix:

$$\mathbf{L} = \begin{pmatrix} \cos\omega_1 & \cdots & \cos(L+1)\omega_1 \\ & \vdots & \vdots \\ \cos\omega_m & \cdots & \cos(L+1)\omega_m \\ \sin\omega_1 & \cdots & \sin(L+1)\omega_1 \\ & \vdots & \vdots \\ \sin\omega_m & \cdots & \sin(L+1)\omega_m \end{pmatrix} \tag{18}$$

and where $\mathbf{p}(d)$ is the $2m$ prediction vector:

$$\mathbf{p}(d) = \begin{pmatrix} \cos(1-d)\omega_1 \\ \vdots \\ \cos(1-d)\omega_m \\ \sin(1-d)\omega_1 \\ \vdots \\ \sin(1-d)\omega_m \end{pmatrix} \tag{19}$$

First and second plots 510, 520 of Figure 5 depict the magnitude and phase response, respectively, of a fixed filter 230 designed using the approach described above with $d = 0$, $L = 32$, $m = 1$, $\omega_1 = 200\,\text{Hz}$ and a sampling rate of 8000 Hz.

-16-

As shown by dashed lines in Figure 5, the fixed filter 230 provides unity gain and zero phase at the center frequency 200Hz as desired.

As described above, the adaptive algorithm used to update the adaptive filter 235 will cause the adaptive filter 235 to converge toward a delayed and scaled replica of the fixed filter 230. Specifically, for a fixed FIR filter 230 with $d = 0$ (i.e., coefficients c_0), the coefficients of the adaptive filter 235 will converge as follows:

$$\hat{c}(k) \sim Sc_{-D} \qquad\qquad (20)$$

where S and D are the scale factor and time delay, respectively, caused by the physical separation of the sensors 215, 225. Thus, the present invention teaches that estimates of the scale factor S and the time delay D can be calculated from the vector relation given by equation (20). For example, if $\hat{S}_k$ denotes an estimate of S based on the coefficients $\hat{c}(k)$ of the adaptive filter 235, then the estimate $\hat{S}_k$ can be computed from equation (20) as follows:

$$\hat{S}_k c_{-D} = \hat{c}(k) \qquad\qquad (21)$$

Furthermore, S can be estimated without prior knowledge of D. To see this, first note that equation (21) is a vector equality in which both c_{-D} and $\hat{c}(k)$ are vectors of size $L + 1$. Then, from equation (17), it follows that $Lc_{-D} = p(-D)$. Thus, multiplication of both sides of equation (21) from the left with the $2m$ x $(L + 1)$ matrix L defined in equation (18), gives the following result:

$$\hat{S}_k p(-D) = L\hat{c}(k) \qquad\qquad (22)$$

where both sides of equation (22) are $2m$-vectors. Then, given that $p(D)^T p(D) = m$ (see equation (19)), equation (22) can be re-written as follows:

$$\hat{S}_k^2 m = (L\hat{c}(k))^T L\hat{c}(k) \qquad\qquad (23)$$

-17-

From equation (23), and from the fact that $S > 0$, the estimate $\hat{S}_k$ of the scale factor S at time instant k can be computed as:

$$\hat{S}_k = \frac{\sqrt{(\mathbf{L}\hat{c}(k))^T \mathbf{L}\hat{c}(k)}}{m} \qquad (24)$$

Given an estimate $\hat{S}_k$ of the scale factor S, an estimate D_k of the time delay D can be computed using a least squares fit as follows:

$$\hat{D}_k = \underset{D}{\arg\min} \; \|\mathbf{L}\hat{c}(k) - \hat{S}_k \mathbf{p}(-D)\|^2 \qquad (25)$$

5   Equivalently, an estimate D_k of the time delay D can be computed as:

$$\hat{D}_k = \underset{\bar{D}}{\arg\min} \; \|\mathbf{L}\hat{c}(k) - \hat{S}_k \mathbf{p}(\bar{D}-\Delta)\|^2 \qquad (26)$$

Advantageously, the estimate D_k can be computed iteratively in practice. Note that the delay gradient $d\mathbf{p}(D)/dD$ follows readily from equation (19).

Thus, the present invention teaches that estimates of the scale factor S and the time delay D can be computed in a straightforward fashion. Advantageously,

10   each of the above described computations can be carried out using well known digital signal processing components. Due to the consistent prior information provided by the fixed filter 230, the estimates will be valid even in the presence of background noise.

The system can be further enhanced by the addition of an activity detector

15   which ensures proper system performance even when all signal sources are inactive. For example, when neither source 210, 220 is active, the signals $x_1(k)$ and $x_2(k)$ received at the sensors 215, 225 will comprise uncorrelated noise only.

In such case, the adaptive filter coefficients $\hat{c}(k)$ will converge toward the null vector, meaning that the scale factor estimate $\hat{S}_k$ will tend toward zero while the time delay estimate $\hat{D}_k$ may take any value. To prevent this, the estimates $\hat{S}_k$, $\hat{D}_k$ can be explicitly set to appropriate values when an activity detector senses the absence of signals of interest.

An exemplary activity detector compares an estimate of the filter noise gain to a predetermined threshold (i.e., an expected noise gain value). An appropriate threshold can be derived from equation (12) as follows:

$$
\begin{aligned}
NG &= c_D^T c_D \\
&= (L^T(LL^T)^{-1}p(D))^T L^T(LL^T)^{-1}p(D) \\
&= p(D)^T(LL^T)^{-1}LL^T(LL^T)^{-1}p(D) \\
&= p(D)^T(LL^T)^{-1}p(D)
\end{aligned}
\qquad (27)
$$

In operation, the activity detector computes an estimate NG of the filter noise gain as a sum of squares of the adaptive filter taps (i.e., $NG = \hat{c}(k)^T\hat{c}(k)$). If the estimate NG is much smaller than the predetermined threshold, then the delay estimate $\hat{D}_k$ is set to zero, and the scale factor estimate $\hat{S}_k$ is set to unity to ensure proper system operation. Note that since the value of the noise gain NG is independent of the delay D, it is sufficient that a single threshold be stored.

An exemplary system can be implemented using the following pseudocode. Those skilled in the art will appreciate that such pseudocode is readily adapted for implementation using standard digital signal processing components.

```
Scale Factor and Time Delay Estimation Routine
Filtering: compute output from the fixed FIR filter and
the adaptive FIR filter (k denotes the running time
index).
Y1 = y1(k:-1:k-L);
Y2 = y2(k:-1:k-L);
```

-19-

```
ylhat(k)  = Y2' * C;
ylfil(k)  = Y1' * CO;
err(k)    = ylfil(k) - ylhat(k);
```

Energy calculations and gain control: A simple gain
control scheme is used in order to set the gain μ to
zero if there is low energy in the inputs. The
instantaneous energy is compared with a long time
average.

```
emom(k)  = sum(ylhat(k:-1:k-L).^2);
eave(k)  = 0.999 * eave(k-1) + 0.001 * emom(k);
if (emom(k) >.5*eave(k))
g(k)  = mu;
else
g(k)  = 0;
end
```

N-LMS update: Update of the adaptive filter
coefficients using the N-LMS algorithm.

```
C = C + g(k) * Y2 * err(k) / ( (Y2'*Y2) + 0.01);
```

Update of estimates of S and $\bar{D}$: The scaling estimate is
smoothed by a first order recursion, while $\bar{D}$ is
estimated by an iterative gradient method. delta
denotes the fixed time delay in channel 1.

```
LLC  =       LL*C;
PPD  =       [cos(warr*(1-Dhat+delta)); sin(warr*(1-
Dhat+delta))];
DPD  =       [sin(warr*(1-Dhat+delta)); -cos(warr*(1-
Dhat+delta))];
Shat =       (1-mu)*Shat + mu*sqrt((LLC'*LLC)/m);
```

-19a-

Dhat =   Dhat + mu*DPD'*(LLC - Shat * PPD);

Activity detector: If the sum square of estimated
filter taps are 20dB below the sum square of the
expected filter taps, the gain is forced to unity and
the delay estimate towards zero.

```
eC = C'*C;
if (eC < 0.01 * eC0)
Shat = 1;
Dhat = 0;
end
```

The estimates of *S* and *D* are further smoothed by a
first order running mean:

```
Sh(k)=rho*Sh(k-1) + (1-rho)*Shat;
Dh(k)=rho*Dh(k-1) + (1-rho)*Dhat;
```

-20-

```
Sh(k)=rho*Sh(k-1)  +  (1-rho)*Shat;
Dh(k)=rho*Dh(k-1)  +  (1-rho)*Dhat;
```

To further illustrate operation of the exemplary embodiments, a numerical example using the pseudocode above is provided. In the example, an acoustic scenario is considered in which the sensors are presumed to be microphones and the sources are presumed to be human speakers or loudspeakers transmitting human speech. As noted above, such a scenario can arise in the context of hands-free mobile telephony used in an automobile environment. Though the example is restricted to two sensors and two sources, those skilled in the art will appreciate that the approach can be applied using an arbitrary number of sources and sensors.

With a distance of 0.5 meters between the first source 210 and the first sensor 215, and with the first sensor 215 treated as a reference sensor, the actual time delay at the second sensor 225 with respect to the first source 210 is $\bar{D} =$ 2.25 samples for a sampling rate of 8kHz. Using the same assumptions, the actual time delay at the second sensor 225 with respect to the second source 220 is $\bar{D} =$ -8.75 samples. These assumptions are reasonable, for example, for a car compartment with a mobile phone (containing the second sensor 225) placed in its cradle near the passenger (the second source 220) and an extra stick-on microphone (the first sensor 215) placed on the sun visor in front of the driver (the first source 210).

In such a car compartment, a rather severe background noise is typically present (e.g., from an AC-fan, the car engine, the road, the wind etc.). For purposes of the numeric example, the sensitivities of the microphones in different directions are assumed to be as shown in Table 1.

-21-

TABLE 1 - MICROPHONE SENSITIVITY TO DIFFUSE BACKGROUND NOISE AND SIGNAL SOURCES IN DIFFERENT POSITIONS.

Signal Source	First Sensor 215 (e.g., sun-visor microphone)	Second Sensor 225 e.g., built-in microphone)
Diffuse Background Noise	0dB	0dB
First Source 210 (e.g., driver)	+3dB	0dB
Second Source 220 (e.g., passenger)	-10dB	0dB

Additionally, synthetic two-channel measurements were created with a male speaker at the location of the first source 210 and a female speaker at the location of the second source 220. The files were concatenated such that there was no speaker activity the first second, then the male speaker was active for 7s, then there was no activity for 3s, and then the female speaker was active for 10s. The signal to noise ratio with respect to the second sensor 225 was 8dB for the male speaker and 7dB for the female speaker (measured over the total period during which they were active). The speech signal detected at the first and second sensors 215, 225 are depicted in first and second plots 610, 620, respectively, of Figure 6.

Additive background noise was modeled as white Gaussian noise. The noise signals detected at the first and second sensors 215, 225 are depicted in third and fourth plots 630, 620, respectively, of Figure 6. The combined speech and noise signals measured at the first and second sensors 215, 225 are depicted in fifth and sixth plots 650, 660, respectively of Figure 6.

In the simulation, the parameters used were: $L = 32$, $\Delta = 10$, $\omega = 2\pi200/8000$, $m = 1$, $\mu = 0.01$ and rho $= 0.99$. The results are shown in Figure 7. Specifically, the delay estimate D_k is depicted in a first plot 710, and the scale

-22-

factor estimate $\hat{S}_k$ is depicted in a second plot 720. In both plots 710, 720, every 50-th sample is displayed. Horizontal dashed lines indicate delays of -3, 0, and 9 samples as well as gains of -10dB, 0dB, and 3dB. As shown, the system properly provides scale factor and time delay estimates, respectively, of 0db and

5   approximately -3 samples when the driver is speaking and -10db and approximately 9 samples when the passenger is speaking. Additionally, the activity detector properly sets the scale factor and time delay estimates, respectively, to 0db and 0 samples during the period when both the driver and the passenger are silent.

10   Although the embodiments have been described in the context of causal filtering (i.e., $\Delta > 0$), the teachings of the present invention are equally applicable in the context of non-causal filtering. Specifically, for $\Delta = 0$ the adaptive scheme comprises an adaptive block that can serve as a signal smoother, a backward predictor ($D < 0$) and/or a forward predictor ($D > 0$). Thus, it is

15   not necessary that a fixed delay be added in the signal flow (e.g., via delay block 410), and an adaptive scheme with minimum inherent delay can be realized. Such a property can be of substantial practical relevance in many real-time applications. However, because the quality of the estimates will be somewhat less accurate with a non-causal approach (and because narrower pass-bands are required for the fixed

20   FIR filter 230), the precise value of Δ can be set based upon system design considerations. For example, Δ can be set to cover "most situations" and not "all possible situations" since the system will provide reasonable results even in rare extreme situations.

Those skilled in the art will appreciate that the present invention is not

25   limited to the specific exemplary embodiments which have been described herein for purposes of illustration. The scope of the invention, therefore, is defined by

-23-

the claims which are appended hereto, rather than the foregoing description, and all equivalents which are consistent with the meaning of the claims are intended to be embraced therein.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1.   A signal processing device, including:

a first signal sensor;

a first filter having an input coupled to an output of said first sensor;

a second signal sensor;

a second filter having an input coupled to an output of said second sensor and having an adjustable filtering characteristic;

a summing device having a first input coupled to an output of said first filter and a second input coupled to an output of said second filter, wherein said summing device outputs a difference between the first and second inputs, and wherein the adjustable filtering characteristic of said second filter is adjusted in dependence upon the output of said summing device; and

a processor for computing an estimate of at least one parameter describing a relationship between said first and second sensors, wherein the at least one parameter indicates at least one of a relative time delay and a relative scale factor between said first and second sensors with respect to a signal source, and wherein the estimate is computed as a function of a filtering characteristic of said first filter and as a function of the adjustable filtering characteristic of said second filter.

2.   A signal processing device according to claim 1, wherein said first filter, said second filter, said summing device and said processor are implemented using a digital signal processor (DSP) integrated circuit (IC).

3.   A signal processing device as claimed in claim 1, wherein said first filter, said second filter, said summing device and said processor are implemented using an application specific integrated circuit (ASIC).

4.   A signal processing device as claimed in claim 1, wherein said signal processing device is a telephone and wherein said first and second sensors are microphones.

5.   A signal processing device as claimed in claim 1, wherein said signal processing device is a radio transceiver and wherein said first and second sensors are antenna elements.

6.   A signal processing device as claimed in any one of claims 1 to 5, wherein said first filter is a finite impulse response (FIR) filter having a fixed filtering characteristic.

7.   A signal processing device as claimed in any one of claims 1 to 5, wherein the filtering characteristic of said first filter includes at least one passband providing unity gain and zero phase delay at a center frequency of the passband.

8.   A signal processing device as claimed in any one of claims 1 to 5, wherein the filtering characteristic of said first filter includes coefficients which are set to minimize a noise gain of said first filter.

9.   A signal processing device as claimed in any one of claims 1 to 5, wherein the filtering characteristic of said filter includes coefficients which are adjusted to optimize a signal-to-noise ratio of said first fiter.

10.   A signal processing device as claimed in any one of claims 1 to 5, wherein the adjustable filtering characteristic of said second filter is adjusted using a Normalized Least Mean Squares (NLMS) algorithm.

11.   A signal processing device as claimed in any one of claims 1 to 5, wherein the adjustable filtering characteristic of said second filter if adjusted using a Least Mean Squares (LMS) algorithm.

12.   A signal processing device as claimed in any one of claims 1 to 5, wherein the adjustable filtering characteristic of said second filter is adjusted using a Recursive Least Squares (RSL) algorithm.
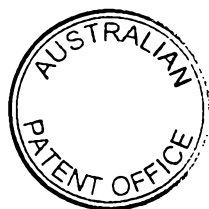
13.   A signal processing device as claimed in claim 1, further including a beamformer for shaping a beam pattern provided by said first and second sensors in dependence upon the estimates of the relative time delay and the relative scale factor between said first and second sensors.

14.   A signal processing device as claimed in claim 13, wherein the beam pattern includes a spacial beam aimed toward a particular signal source.

15.   A signal processing device as claimed in claim 13, wherein the beam pattern includes a spacial notch aimed against a particular signal source.

16.   A signal processing device according to claim 1, wherein said processor selects a signal detected by a particular one of said first and second sensors for transmission in dependence upon the estimates of the relative time delay and the relative scale factor between said first and second sensors.

17.   A signal processing device as claimed in claim 1, further including at least one additional sensor and at lest one additional filter having and adjustable filtering characteristic,

wherein the adjustable filtering characteristic of said additional filter is adjusted in dependence upon a difference between the output of said first filter and an output of said additional filter, and

wherein said processor computes an estimate of at least one parameter relating to said first sensor and said additional sensor in dependence upon the filtering characteristic of said first filter and the adjustable filtering characteristic of said additional filter.

18.   A signal processing device as claimed in claim 17, wherein said processor computes estimates of a relative time delay and a relative scale factor between said first sensor and said additional sensor with respect to a signal source.

19.   A signal processing device as claimed in claim 1, further including an activity detector for detecting when a signal source of interest is active, wherein said processor sets the estimates of the relative time delay and the relative scale factor to predetermined values when said activity detector indicates that no signal source of interest is active.

20.   A signal processing device as claimed in claim 19, wherein said processor sets the estimate of the relative time delay to zero and the estimate of the scale factor to one when said activity detector indicates that no signal source of interest is active.
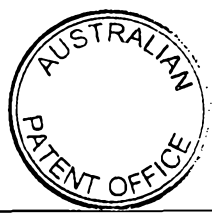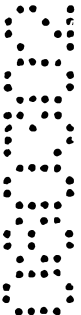
21.   A signal processing device as claimed in claim 1, further including
a fixed-delay block positioned in a signal flow path corresponding to said first
sensor.

22.   A signal processing device as claimed in claim 1,
wherein filtering characteristics of said first and second filters each
include a number L of filtering coefficients,
wherein the filtering characteristic of said first filter includes a
number m of passbands, each passband ℓ, $\ell \in (1,m)$, having a center frequency ω_ℓ,
and
wherein an estimate D_k of a relative time delay D between said first
and second sensors and an estimate $\hat{S}_k$ of a relative scale factor S between said first
and second sensors, at a time instant k, are computed based upon the adjustable
filtering characteristic $\hat{c}(k)$ of said second filter, a matrix L, and a prediction
vector $p(D)$, as:

$$\hat{S}_k = \frac{\sqrt{(L\hat{c}(k))^T L\hat{c}(k)}}{m}$$

$$\hat{D}_k = \begin{matrix} \arg\min \\ D \end{matrix} \| L\hat{c}(k) - \hat{S}_k p(-D) \|^2$$

where the matrix L and the prediction vector $p(D)$ are computed as:

$$L = \begin{pmatrix} \cos\omega_1 & \cdots & \cos(L+1)\omega_1 \\ \vdots & & \vdots \\ \cos\omega_m & \cdots & \cos(L+1)\omega_m \\ \sin\omega_1 & \cdots & \sin(L+1)\omega_1 \\ \vdots & & \vdots \\ \sin\omega_m & \cdots & \sin(L+1)\omega_m \end{pmatrix}$$
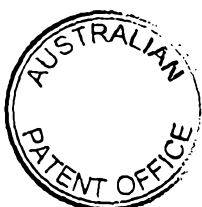
$$p(D) = \begin{pmatrix} \cos(1-D)\,\omega_1 \\ \vdots \\ \cos(1-D)\,\omega_m \\ \sin(1-D)\,\omega_1 \\ \vdots \\ \sin(1-D)\,\omega_m \end{pmatrix} .$$

23.   A method for processing signals, including the steps of:

detecting a first signal using a first signal sensor;

filtering the first signal using a first filter to provide a first filtered

signal;

detecting a second signal using a second signal sensor;

filtering the second signal using a second filter to provide a second filtered

signal;

computing a difference between the first filtered signal and the second

filtered signal;

adjusting a filtering characteristic of the second filter in dependence upon

the difference obtained in said step of computing; and

estimating at least one parameter as a function of a filtering characteristic

of said first filter and as a function of the filtering characteristic of said second

filter, wherein the at least one parameter indicates at least one of a relative time

delay and a relative scale factor between said first and second sensors with

respect to a signal source.

24.   A method according to claim 23, wherein the filtering characteristic of the

first filter includes at least one passband providing unity gain and zero phase

delay at a center frequency of the passband.

25. A method as claimed in claim 23, wherein the filtering characteristic of said second filter is adjusted using a Normalized Least Mean Squares (NLMS) algorithm.

26. A method as claimed in claim 23, further including a step of shaping a beam pattern provided by said first and second sensors in dependence upon the relative time delay and the relative scale factor obtained as a result of said step of estimating.

27. A method as claimed in claim 23, further including a step of selecting and transmitting a signal detected by a particular one of said first and second sensors in dependence upon the relative time delay and the relative scale factor obtained as a result of said step of estimating.

28. A method as claimed in claim 23, further including the steps of detecting whether a signal source of interest is active and setting relative time delay and scale factor estimates to predetermined values when said step of detecting indicates that no source of interest is active.

29. A device as claimed in claim 1 substantially as herein described with reference to Figures 2 to 7 of the accompanying drawings.

30. A method as claimed in claim 23 substantially as herein described with reference to Figures 2 to 7 of the accompanying drawings.

DATED this 6th day of March 2002
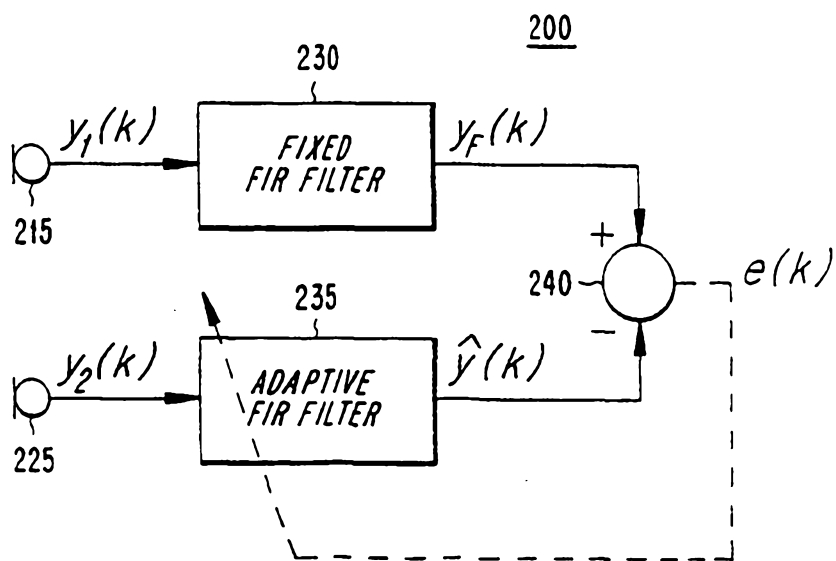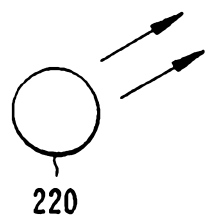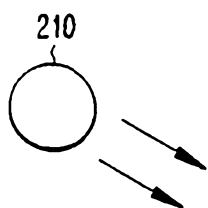TELEFONAKTIEBOLAGET L M ERICSSON

WATERMARK PATENT & TRADE MARK ATTORNEYS
290 BURWOOD ROAD
HAWTHORN VICTORIA 3122
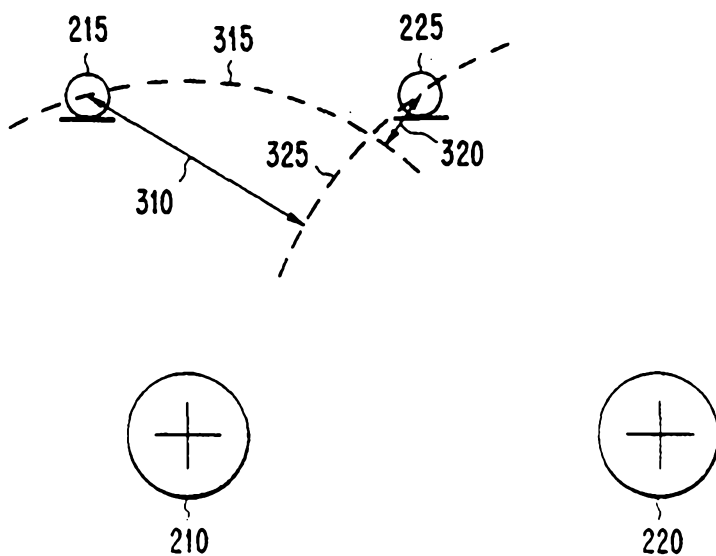AUSTRALIA
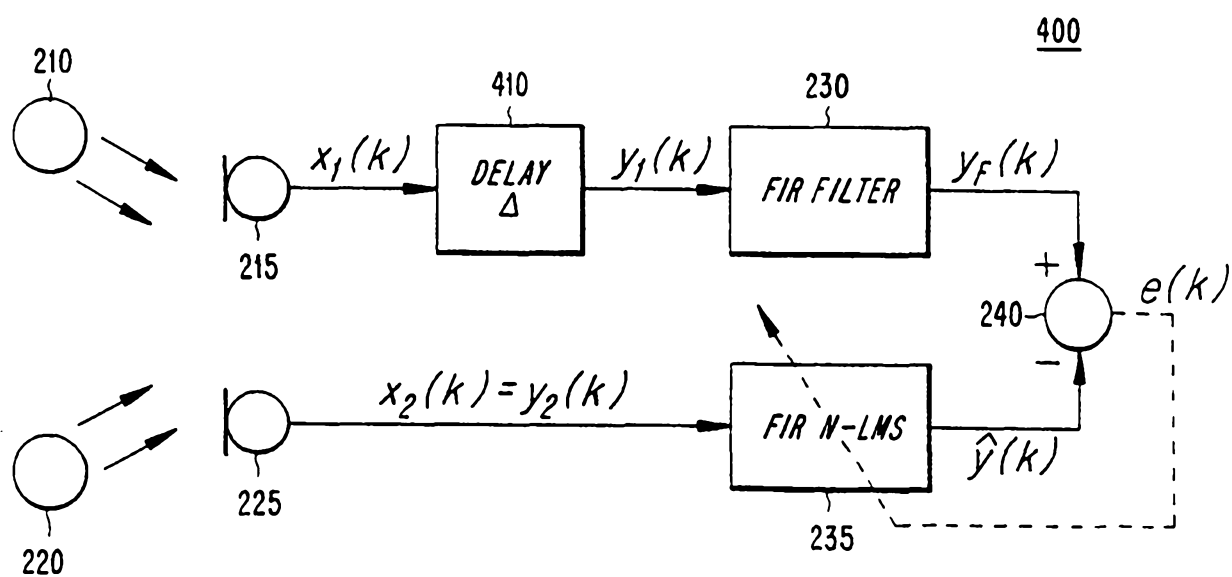
RCS/AJL/MBL
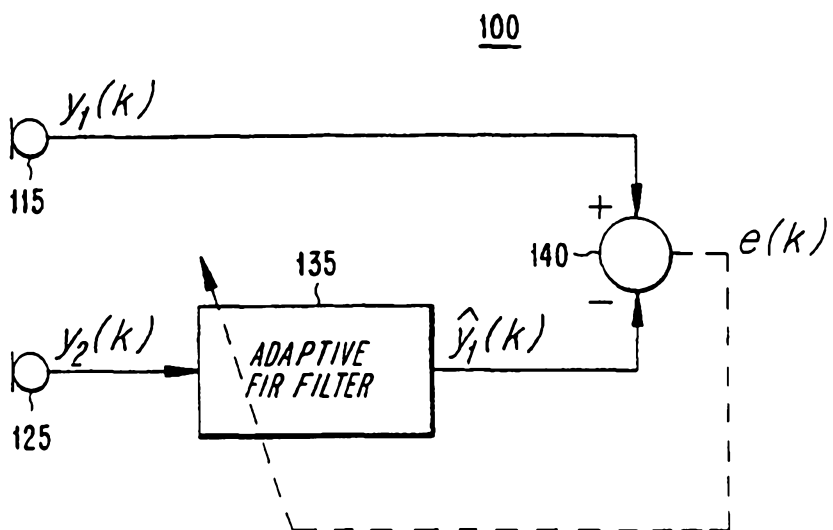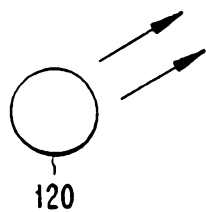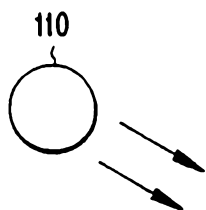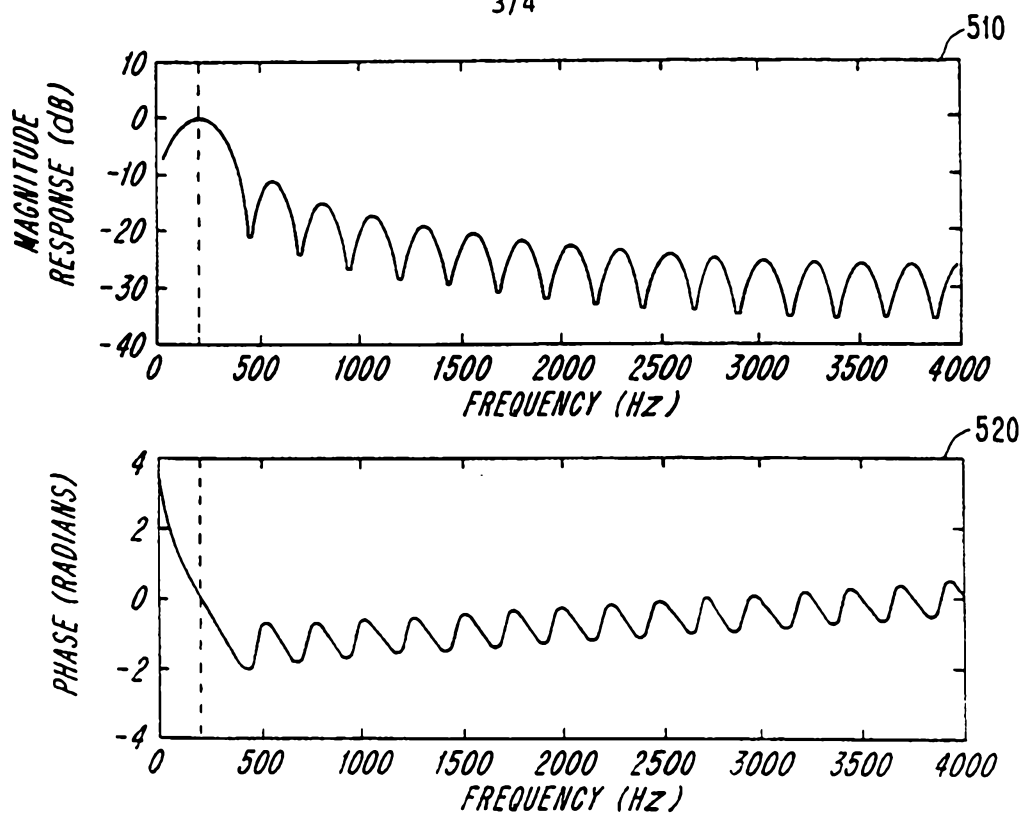
FIG. 1

(PRIOR ART)

FIG. 2

FIG. 3

FIG. 4

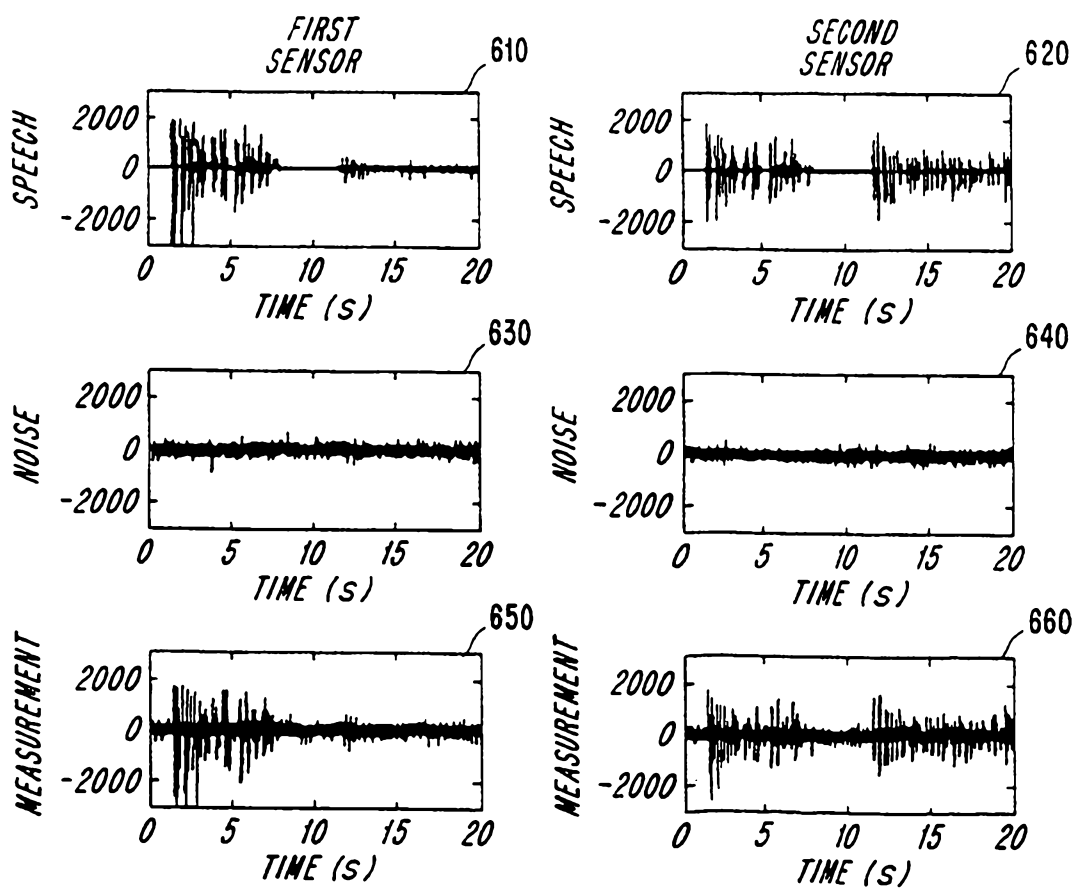FIG. 5

FIG. 6

FIG. 7