

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
10 January 2002 (10.01.2002)

PCT

(10) International Publication Number
WO 02/03749 A2

(51) International Patent Classification⁷: **H04R**

(21) International Application Number: PCT/US01/18416

(22) International Filing Date: 5 June 2001 (05.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
09/593,728 13 June 2000 (13.06.2000) US

(71) Applicant: **GN RESOUND CORPORATION** [US/US];
220 Saginaw Drive, Seaport Centre, Redwood City, CA
94063 (US).

(72) Inventors: **LUO, Fa-Long**; 1536 Oxford Street, #2, Red-
wood City, CA 94061 (US). **YANG, Jun**; 1536 Oxford

Street, #2, Redwood City, CA 94061 (US). **EDWARDS, Brent**; 250 Point Lobos #203, San Francisco, CA 94121 (US). **MICHAEL, Nick**; 39 McCoppin Street, San Francisco, CA 94103 (US).

(74) Agent: **KREBS, Robert E.**; Burns, Doane, Swecker & Mathis, LLP, P.O. Box 1404, Alexandria, VA 22313 (US).

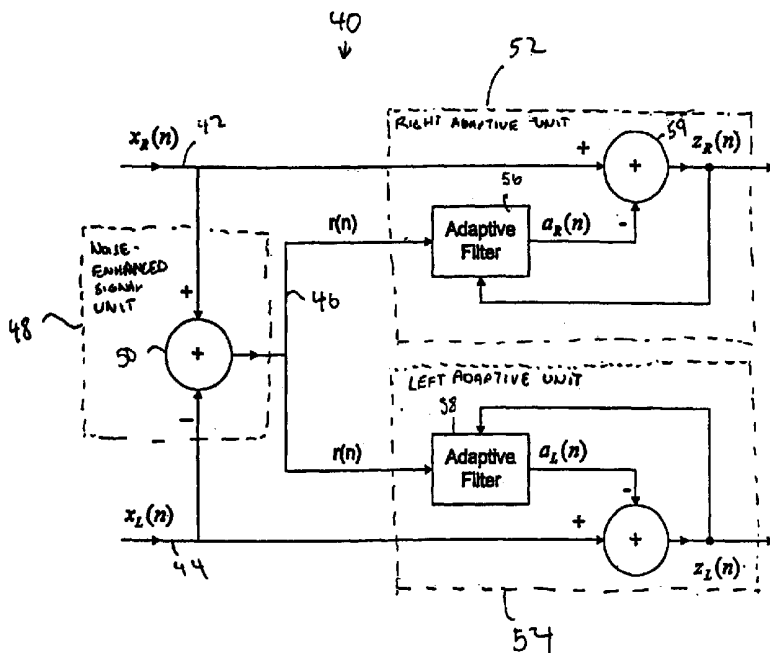
(84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR).

Published:

— without international search report and to be republished upon receipt of that report

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

(54) Title: ADAPTIVE MICROPHONE ARRAY SYSTEM WITH PRESERVING BINAURAL CUES



(57) Abstract: A microphone system using left and right microphones produces good binaural cues as well as noise reduction by using two adaptive filters. A noise-enhanced reference signal is sent to two different adaptive filters. The output of each adaptive filter is combined with the left or right microphone signal, respectively, to produce left and right output signals. These left and right output signals are used as the error signals to modify the coefficients of the adaptive filters. By using two adaptive filters, each of the adaptive filters is able to independently operate helping to preserve the binaural cues.

WO 02/03749 A2

ADAPTIVE MICROPHONE ARRAY SYSTEM WITH PRESERVING BINAURAL CUES

Background of the Invention

The present invention relates to adaptive microphone array systems.

5 The combination of multiple-microphone-based spacial processing with binaural listening is receiving increasing attention in many application fields such as hearing aids because this combination should be able to provide both the spacial filtering benefits of a microphone array and the natural benefits of binaural listening for sound-location ability and speech intelligibility. Typically, a
10 microphone is positioned on each side of a user's head. Ideally, the system is designed such that the two microphones can be used to improve the signal-to-noise ratio as well as to maintain the binaural cues.

One proposed system is described in the reference, *Microphone-Array Hearing Aids With Binaural Output — Part I: Fixed-Processing Systems*,
15 and *Part II: A Two-Microphone Adaptive System*, IEEE Transactions on Speech and Audio Processing, Vol. 5, No. 6 (Nov. 1997) pp 529-551 and shown in Fig. 1. In this example, $X_R(n)$ is the microphone signal received at the right ear, and $X_L(n)$ is the microphone signal received at the left ear. A low-pass filter and a high-pass filter with cut-off frequency f_c are used at each channel. The outputs
20 $f_R(n)$ and $f_L(n)$ of the two high-pass filters are sent to an adaptive processor 22 whose output is $Y(n)$. The output of the two low-pass filters 24 and 26 are delayed in delays 28 and 30, and combined with the output $Y(n)$ of the adaptive processor 22 to provide binaural outputs $Z_R(n)$ and $Z_L(n)$. In this prior-art scheme, the combination of the adaptive array processing with binaural listening is
25 accomplished by dividing the frequency spectrum. The low-pass portion of the frequency spectrum is devoted to binaural processing, and the high-pass part of the spectrum is devoted to adaptive array processing. The binaural cues of signal

-2-

parts with higher frequencies than the cut-off frequency will be lost in the system. Likewise, the benefits of the adaptive array processing for lower frequencies than the cut-off frequency also will be lost. Furthermore, two low-pass filters and two high-pass filters are required. More importantly, appropriate equalization
5 processing between the adaptive array processing and the outputs of the low-pass filters are required so as to avoid unexpected artifacts that result from a simple cutoff of the spectrum and simple summation. These problems make this prior system complicated and fail as a practical solution in achieving maximum array processing benefits and binaural listening benefits.

10 It is desired to have an improved adaptive array microphone system that maintains binaural cues and provides spacial filtering benefits.

Summary of the Present Invention

The present invention comprises a microphone system using two adaptive filters, each receiving the same reference signal derived from two ear microphones
15 but having different primary and error (filter adjustment) signals. The primary signals are preferably from the right and left microphone output signals, respectively.

One embodiment of the present invention comprises an apparatus including a noise-enhanced data producing unit receiving left- and right-ear microphone data and producing noise-enhanced data. A right adaptive unit receives the right
20 microphone data and the noise-enhanced data, and produces a reduced-noise right data output. The right adaptive unit includes a first adaptive filter receiving the noise-enhanced data. A left adaptive unit receives the left microphone data and the noise-enhanced data. The left adaptive unit produces reduced-noise left data
25 output. The left adaptive unit includes a second adaptive filter receiving the noise-enhanced data.

-3-

Another embodiment of the present invention is a method including the steps of: calculating the noise-enhanced data from the left and right microphone data; adaptively filtering the noise-enhanced data to produce first filtered noise-enhanced data; combining the first filtered noise-enhanced data with right
5 microphone data to produce a reduced-noise right data output; adaptively filtering the noise-enhanced data to produce second filtered noise-enhanced data; and combining the second filtered noise-enhanced data with left microphone data to produce a reduced-noise left data output.

Brief Description of the Drawings

10 Fig. 1 is a diagram of a prior-art microphone system.

Fig. 2 is a diagram of a microphone system of one embodiment of the present invention.

Fig. 3 is a diagram of one embodiment of the present invention implementing the system of Fig. 2.

15 Fig. 4 is a flow chart illustrating the operation of one embodiment of the present invention.

Detailed Description of the Preferred Embodiment

Fig. 2 is a diagram illustrating the functional portions of the microphone system of the present invention. $X_R(n)$ 42 comprises data from the right
20 microphone. $X_L(n)$ 44 comprises the data from the left microphone. Data $r(n)$ is derived from these two data signals. In a preferred embodiment, a noise-enhanced signal unit 48 produces an output, $r(n)$, with enhanced noise. In one embodiment, the noise-enhanced signal unit 48 includes a summing unit 50 which subtracts one of the microphone signals from the other. This produces a noise-enhanced signal
25 since sound coming from the front, including presumably the desired speech

-4-

signal, will reach the left and right microphones at nearly the same time, thus forming a null in the output response at the front.

The noise-enhanced signal is applied to the right adaptive unit 52 and left adaptive unit 54. Preferably, the right adaptive unit includes a first adaptive filter 56 receiving the noise-enhance signal $r(n)$. The right adaptive unit 52 preferably receives $X_R(n)$ as a primary signal to adjust the adaptive filter 56. In one embodiment, the unit 52 also includes a summing unit 59, which subtracts the output of the adaptive filter 56 from the right front signal $X_R(n)$ to produce the output $Z_R(n)$. The left adaptive unit also includes a second adaptive filter 58. The second adaptive filter also receives the reference signal $r(n)$ and produces an output $a_L(n)$. The left adaptive unit 54 receives the output $X_L(n)$ as the primary signal to modify the adaptive filter coefficients.

Note that although the first and second adaptive filters receive the same reference signal, the coefficients of the adaptive filters are different since the primary and error signals used to modify the coefficients are different for the first and second adaptive filters.

As described below, the adaptive filters can be of a variety of different types which are known in the art. In one preferred embodiment, the adaptive filters use the same algorithm but, as described above, have different primary signals.

One embodiment of the system of the present invention is described in the following mathematical description. The received signals at the right ear microphone and the left ear microphone are $X_R(n)$ and $X_L(n)$ which consist of the target speech parts $S_R(n)$, $S_L(n)$ and the noise parts $N_R(n)$, $N_L(n)$, respectively, that is

$$X_R(n) = S_R(n) + N_R(n)$$

and

$$X_L(n) = S_L(n) + N_L(n)$$

-5-

The reference signal $r(n) = X_R(n) - X_L(n)$ is sent to two adaptive filters with weights $W_R(n) = [W_{R1}(n), W_{R2}(n), \dots, W_{RN}(n)]^T$ and $W_L(n) = [W_{L1}(n), W_{L2}(n), \dots, W_{LN}(n)]^T$ and these two adaptive filters provide the outputs $a_R(n)$ and $a_L(n)$ as follows, respectively:

$$a_R(n) = \sum_{m=1}^N W_{Rm}(n) r(n-m+1) = W_R^T(n) R(n)$$

5 and

$$a_L(n) = \sum_{m=1}^N W_{Lm}(n) r(n-m+1) = W_L^T(n) R(n)$$

where $R(n) = [r(n), r(n-1), \dots, r(n-N+1)]^T$ and N is the length of two adaptive filters. Note that the length of two filters is here selected to be the same for simplicity but could be different. The primary signals at two adaptive filters are $X_R(n)$ and $X_L(n)$, and two outputs: $Z_R(n)$ and $Z_L(n)$ for the right ear and the left ear, respectively, are obtained by

10

$$Z_R(n) = X_R(n) - a_R(n)$$

and

$$Z_L(n) = X_L(n) - a_L(n)$$

The weights of these two adaptive filters are adjusted so as to minimize the average power of the two outputs, that is,

-6-

$$\min_{W_R(n)} E(|Z_R(n)|^2) = \min_{W_R(n)} E(|X_R(n) - a_R(n)|^2)$$

and

$$\min_{W_L(n)} E(|Z_L(n)|^2) = \min_{W_L(n)} E(|X_L(n) - a_L(n)|^2)$$

In the ideal case $r(n)$ contains only the noise part and the two adaptive filters can provide two outputs $a_R(n)$ and $a_L(n)$ by minimizing the two foregoing equations. Since the reference signal $r(n)$ does not include the signal portion, the adaptive filters will adjust $a_R(n)$ and $a_L(n)$ to remove the noise portion from the primary signals, that is, $a_R(n) \approx N_R(n)$ and $a_L(n) \approx N_L(n)$. As a result, two system outputs $Z_R(n)$ and $Z_L(n)$ will approximate the signal parts $S_R(n)$ and $S_L(n)$, respectively. This means that the above processing not only realizes maximum noise reduction by two adaptive filters but also preserves the binaural cues contained in the target signal parts $S_R(n)$ and $S_L(n)$.

The learning algorithm can be an adaptive algorithm such as the LS, RLS, TLS and LMS algorithms. The LMS algorithm version to update the coefficients of two adaptive filters is

$$W_R(n+1) = W_R(n) + \lambda R(n) Z_R(n)$$

$$W_L(n+1) = W_L(n) + \lambda R(n) Z_L(n)$$

-7-

where λ is a step parameter which is a positive constant less than $\frac{2}{P}$ where P is the power of the input $r(n)$ of these two adaptive filters. For better performance and faster convergence speed, λ can be also time varying as done in the normalized LMS algorithm uses, that is,

$$W_{\mathbf{R}}(n+1) = W_{\mathbf{R}}(n) + \frac{\mu}{\|R(n)\|^2} R(n)Z_{\mathbf{R}}(n)$$

$$W_{\mathbf{L}}(n+1) = W_{\mathbf{L}}(n) + \frac{\mu}{\|R(n)\|^2} R(n)Z_{\mathbf{L}}(n)$$

5 where μ is a positive constant less than 2.

Based on the frame-by-frame processing configuration, a further modified algorithm can be obtained as follows (Denoted by Equation 1 and Equation 2):

$$W_{\mathbf{R}k}(n+1) = W_{\mathbf{R}k}(n) + \frac{\mu}{\|R(n)\|^2} R(n)Z_{\mathbf{R}k}(n)$$

and

$$W_{\mathbf{L}k}(n+1) = W_{\mathbf{L}k}(n) + \frac{\mu}{\|R(n)\|^2} R(n)Z_{\mathbf{L}k}(n)$$

where k represents the k 'th repeating in the same frame.

-8-

In comparison with the prior-art scheme of Fig. 1, the advantages of the present invention include the following: First, no binaural cue of the target speech will be lost because two system outputs $Z_R(n)$ and $Z_L(n)$ will approximate the signal parts $S_R(n)$ and $S_L(n)$ respectively. In the prior scheme, the binaural cues of signals with higher frequencies than the cut-off frequency have been lost. Second, the array processing benefit for the part with frequencies less than the cut-off frequency will be preserved because the frequency spectrum is not divided. Note that in the prior scheme, this benefit also has been lost. Third, in the present invention no low-pass filters and high-pass filters are required, while the prior scheme requires two low-pass filters and two high-pass filters. Consequently, the related equalization processing is not required in the present invention. The only cost involved in implementing the present invention is to include an additional adaptive filter. The two adaptive filters can use the same structure and adaptive algorithm. This property results in a large convenience in hardware implementation because the related assembly code and machine code of two adaptive filters can be shared.

Fig. 3 is a diagram that illustrates an embodiment of the system of the present invention. The system 70 includes a right ear microphone 72 and a left ear microphone 74. The converters 76 and 78 convert the analog signals into digital signals which are sent to the processor 80. In a preferred embodiment, the processor 80 is a digital signal processor. The processor 80 loads the adaptive microphone array program 82 from memory 84. The adaptive microphone array program can implement the functional blocks as shown in Fig. 2. Other programs 86 such as hearing aid algorithms can also be stored in the memory 84 for loading into the processor 80. The output signals of the processor 80 can be sent to speakers if a hearing-aid system is used.

Fig. 4 is a flow chart that illustrates the operation of one embodiment of the present invention. In step 90, an enhanced noise signal is calculated using left and

right microphone samples. This enhanced noise signal can be the reference signal constructed by subtracting one of the microphone samples from the other microphone sample. In step 92, a noise-enhanced signal and the right ear signal are used to produce a noise-reduced right-ear signal. This is preferably done by
5 adaptively filtering the noise-enhanced signal and combining the filtered signal with the right-ear microphone system to produce an output. The output being used as an error signal for the adaptive filter. In step 94, the noise-enhanced signal and the left-ear signal are used to produce a noise-reduced left-ear signal. The output of a second adaptive filter is combined with the left-ear signal, producing the
10 noise-reduced left-ear signal. The noise-reduced left-ear signal is then used as an error signal to affect the coefficients of the adaptive filter. Note that the order of steps 92 and 94 is not important. In step 96 the left and right noise-reduced signals are used to produce coefficients which are used in the adaptive filters used in the producing steps 92 and 94.

15 Any kind of adaptive algorithms, such as LMS-based, LS-based, TLS-based, RLS-based and related algorithms, can be used in this scheme. The weights can also be obtained by directly solving the estimated Wiener-Hopf equations. Moreover, repeated adaptive algorithms like Equations 1 and 2 and adaptive lattice filters can be used in this scheme as well. In one embodiment, the lengths of two
20 adaptive filters are adjustable and can take different values. Also, the step parameters in related adaptive algorithms for two adaptive filters can take different values. Trade-offs between performance and cost (complexity, etc.) in practical applications determine which algorithm is used.

The two adaptive filters in Fig. 2 can be nonlinear filters and can be
25 implemented by some neural networks such as multi-layer perceptron networks, radial basis function networks, high-order neural networks, etc. The corresponding learning algorithms in neural networks such as the back-propagation algorithm can be used for the adaptive filters.

-10-

A matching filter could be added in either the left ear channel or the right ear channel before obtaining the difference signal $r(n)$ so as to compensate for the magnitude mismatch of two-ear microphones. The matching filter can be in either a finite impulse response (FIR) filter or an infinite impulse response (IIR) filter.

5 The matching filter could also be either in a fixed model or in an adaptive model.

In one embodiment, to reduce the target signal cancellation problem existing in most adaptive array processing algorithms, a speech pause detection system is used and the weight update of the two adaptive filters is made during a pause in the speech.

10 Either a directional microphone or an omnidirectional microphone could be used in this invention. Some pre-processing methods could be used to improve the signal-to-noise ratio of two primary signals $X_R(n)$ and $X_L(n)$. These pre-processing methods include that more than one microphone can be used in each ear and then these microphone signals are combined in some way to produce the signals $X_R(n)$ and $X_L(n)$. This pre-processing could be either in a fixed model or in an adaptive
15 model and also either in the spatial domain or in the temporal domain.

It will be appreciated by those of ordinary skill in the art that the invention can be implemented in other specific forms without departing from the spirit or character thereof. The presently disclosed embodiments are therefore considered
20 in all respects to be illustrative and not restrictive. The scope of the invention is illustrated by the appended claims rather than the foregoing description, and all changes that come within the meaning and range of equivalents thereof are intended to be embraced herein.

Claims:

1. An apparatus comprising:
a noise-enhanced data producing unit receiving left and right ear microphone data and producing noise-enhanced data;
a right adaptive unit receiving the right microphone data and the noise-enhanced data, the right adaptive unit producing reduced-noise right data output, the right adaptive unit including a first adaptive filter receiving the noise-enhanced data; and
a left adaptive unit receiving the left microphone data and the noise-enhanced data, the left adaptive unit producing reduced-noise left data output, the left adaptive unit including a second adaptive filter receiving the noise-enhanced data.
2. The apparatus of Claim 1 wherein the first adaptive filter uses the reduced-noise right data output as an error signal to modify the coefficients of the first adaptive filter.
3. The apparatus of Claim 1 wherein the second adaptive filter uses the reduced-noise left data output as an error signal to modify the coefficients of the second adaptive filter.
4. The apparatus of Claim 1 wherein the apparatus is implemented as a software program on a processor-based system.
5. The apparatus of Claim 1 wherein the noise-enhanced-data-producing unit comprises a summer to subtract one of the ear's microphone data from the other ear's microphone data.

-12-

6. The apparatus of Claim 1 wherein the right adaptive unit includes a summer unit to subtract the output of the first adaptive filter from the right microphone data to produce the reduced-noise right-data output.

7. The apparatus of Claim 1 wherein the left adaptive unit includes a summer unit to subtract the output of the second adaptive filter from the left microphone data to produce the reduced-noise left-data output.

8. A method comprising:
calculating noise- enhanced data from left and right microphone data;
adaptive filtering the noise-enhanced data to produce first filtered noise-enhanced data;
combining the first filtered noise- enhanced data with right microphone data to produce a reduced noise right data output;
adaptive filtering the noise-enhanced data to produce second filtered noise-enhanced data; and
combining the second filtered noise- enhanced data with left microphone data to produce a reduced noise left data output.

9. The method of Claim 8 wherein the method is implemented on a processor.

10. The method of Claim 8 wherein the noise-enhanced data-calculating step comprises subtracting one of the microphone data from the other microphone data.

-13-

11. The method of Claim 8 wherein the first adaptive filtering step uses the reduced-noise right data output as an error signal to modify the coefficients of the first adaptive filter.

12. The method of Claim 8 wherein the second adaptive filtering step uses the reduced-noise left data output as an error signal to modify the coefficients of the second adaptive filter.

13. The method of Claim 8 wherein the combining step comprises subtracting the filtered noise-enhanced data from the microphone data.

14. A computer-readable medium containing a program which executes the following procedure:

calculating noise-enhanced data from left and right microphone data;

adaptive filtering the noise-enhanced data to produce first filtered noise-enhanced data;

combining the first filtered noise-enhanced data with right microphone data to produce a reduced noise right data output;

adaptive filtering the noise-enhanced data to produce second filtered noise-enhanced data; and

combining the second filtered noise-enhanced data with left microphone data to produce a reduced noise left data output.

15. The computer-readable medium of Claim 14 wherein the method is implemented on a processor.

-14-

16. The computer-readable medium of Claim 14 wherein the noise-enhanced data-calculating step comprises subtracting one of the microphone data from the other microphone data.

17. The computer-readable medium of Claim 14 wherein the adaptive filtering step uses the reduced-noise data output as an error signal to modify the coefficients of the adaptive filter.

18. The computer-readable medium of Claim 14 wherein the combining step comprises subtracting the filtered noise-enhanced data from the microphone data.

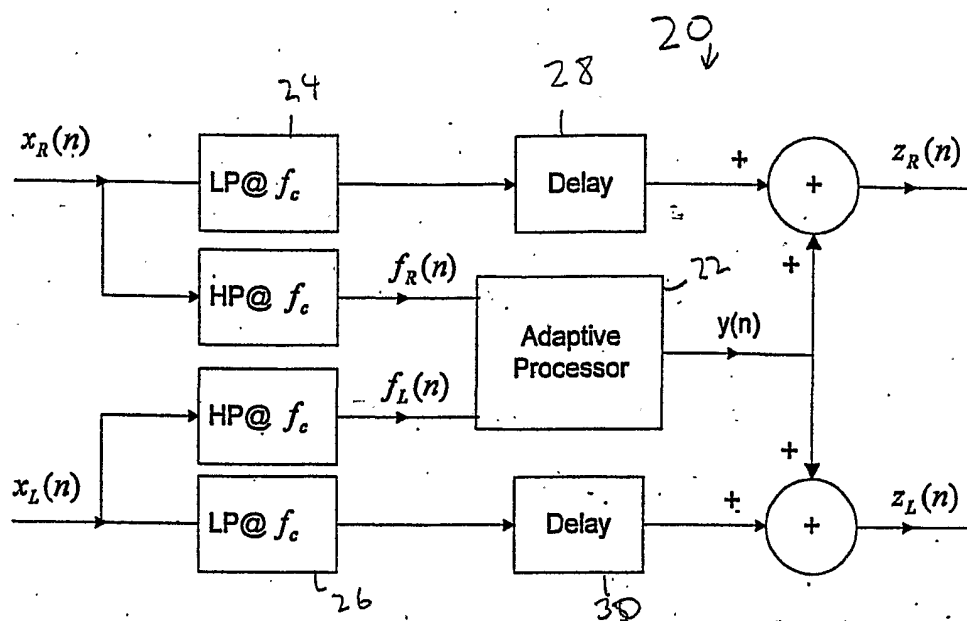
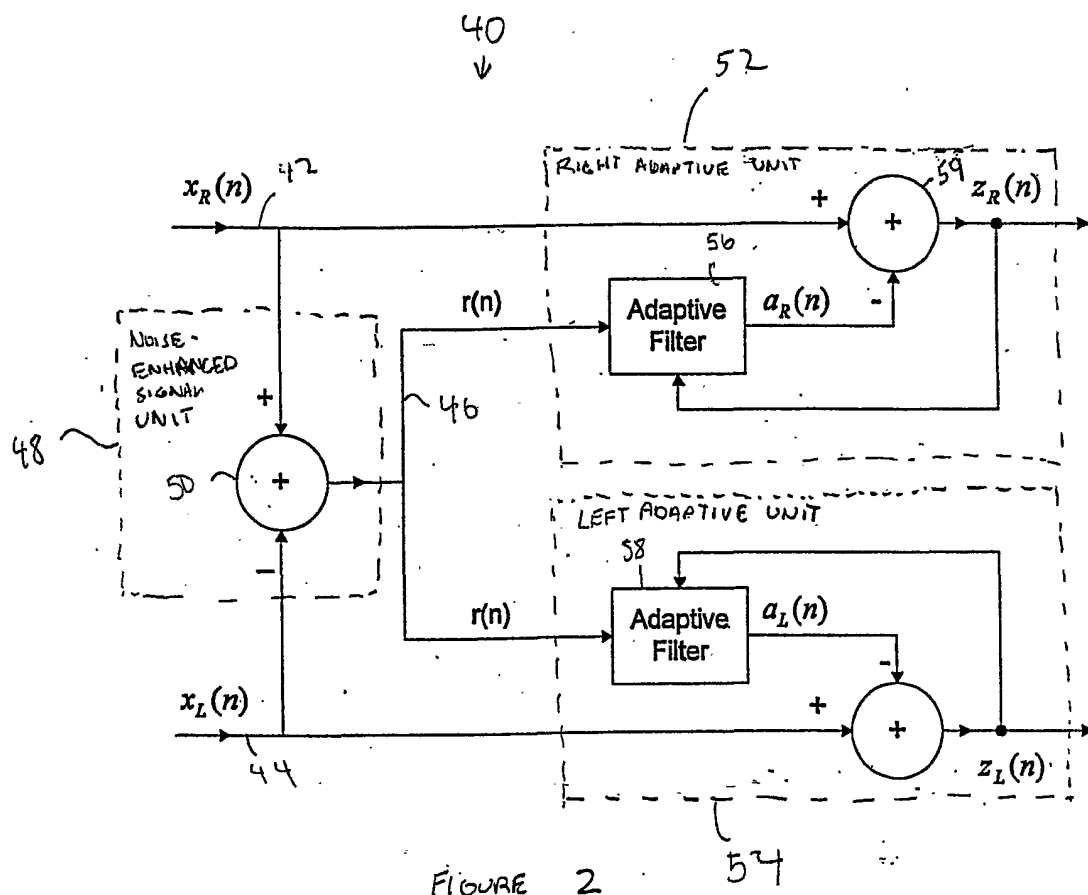


FIGURE 1
(PRIOR ART)



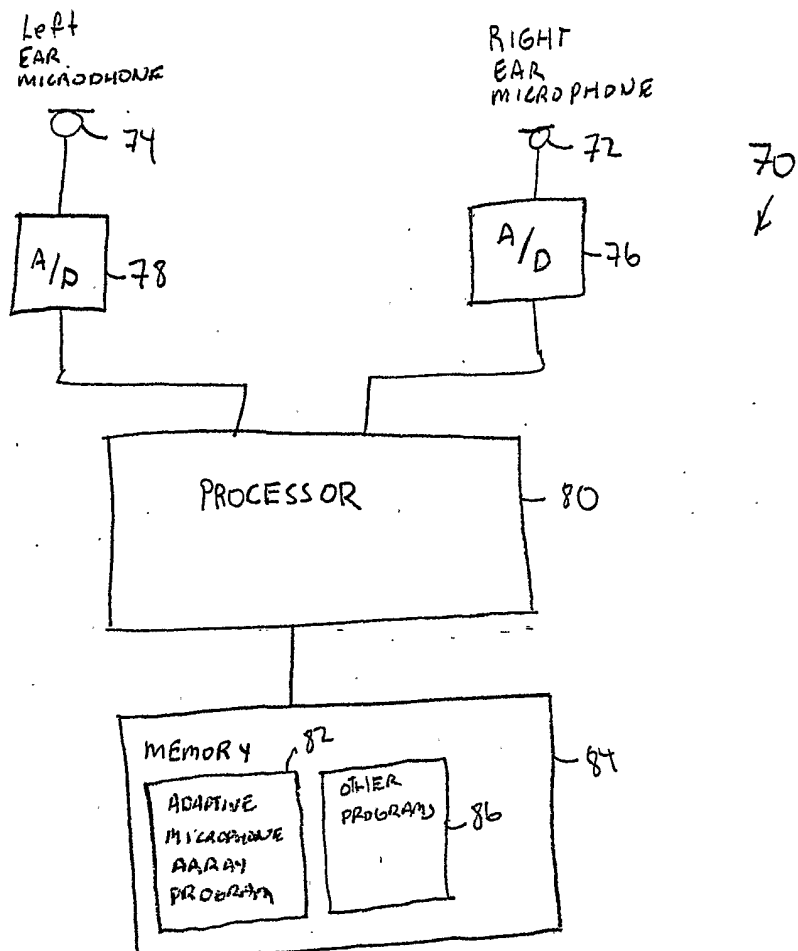


FIGURE 3

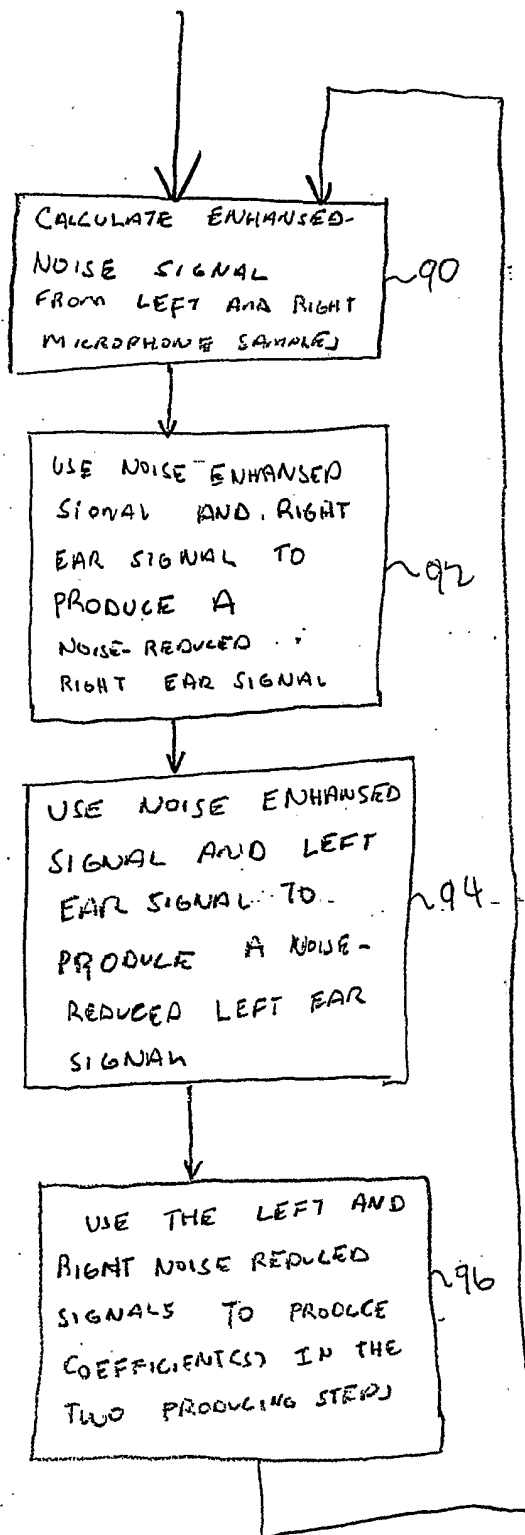


FIGURE 4