

AN ADAPTIVE DUAL FILTER ECHO CANCELLATION METHOD

TECHNICAL FIELD

The present invention relates to an adaptive dual filter echo cancellation method and a method for determining a filter quality measure that is used in said echo cancellation method.

BACKGROUND OF THE INVENTION

Echo is a problem related to the perceived speech quality in telephony systems with long delays, e.g. telephony over long distances or telephony systems using long processing delays, like digital cellular systems. The echo arises in the four-to-two wire conversion in the PSTN/subscriber interface. To remove this echo, echo cancellers are usually provided in transit exchanges for long distance traffic, and in mobile services switching centers for cellular applications.

Due to the location of the echo canceller it is made adaptive; the same echo canceller is used for many different subscribers in the PSTN. This adaption is necessary not only between different calls, but also during each call, due to the non-fixed nature of the transmission network, e.g. phase slips, three-party calls, etc.

The adaption of the echo canceller needs to be controlled, since it must be inhibited during presence of near end side speech, otherwise the echo path estimate will be degraded. This leads to a conservative strategy with a well protected estimate. However, the adaption strategy cannot be too conservative, since this will degrade the performance of the echo canceller when a fast re-adaption is necessary due to a change in the echo path loop. To overcome the optimization problem, namely fast re-adaption when the echo path changes and stable echo estimate during double-talk, a configuration with two echo path estimates may be used. Echo cancellers using two filters for echo estimation have been

described in [1, 2]. One filter, commonly known as the foreground filter, is non-adaptive and used for obtaining the actual echo canceller output. The other filter, commonly known as the background filter, is continuously updated with some adaptive algorithm, typically a normalized least mean square (NLMS) algorithm. The coefficients from the adaptive background filter are then transferred to the foreground filter whenever the background filter is considered better in some sense.

Since the configuration described in [1, 2] only uses the non-adaptive foreground filter for echo canceller output, it is very important that the adaptive background filter is transferred when it performs better. However, due to problems, partly caused by the conservative algorithm that is used, this may not occur and echo cancellation may be inhibited.

SUMMARY OF THE INVENTION

An object of the present invention is to provide a new method of determining a filter quality measure that may be used in selecting the best filter in a dual filter echo canceller.

This method is characterized by the features of claim 1.

A further object of the present invention is an adaptive dual filter echo cancellation method that is less conservative than the previously known method and avoids the problems of that method.

This echo cancellation method is characterized by the features of claim 3.

BRIEF DESCRIPTION OF THE DRAWINGS

The invention, together with further objects and advantages thereof, may best be understood by making reference to the following description taken together with the accompanying

drawings, in which:

- FIGURE 1 is a block diagram of an echo generating system;
- FIGURE 2 is a block diagram of an echo cancellation system;
- 5 FIGURE 3 is a block diagram of a previously known dual filter echo canceller;
- FIGURE 4 is a block diagram of a dual filter echo canceller operating in accordance with the echo cancellation method of the present invention;
- 10 FIGURE 5 is a flow chart illustrating an embodiment of the dual filter echo cancellation method in accordance with the present invention;
- FIGURE 6 is a preferred embodiment of the dual filter echo cancellation method in accordance with the present invention; and
- 15 FIGURE 7 is another preferred embodiment of the dual filter echo cancellation method in accordance with the present invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

20 Fig. 1 illustrates the echo generating process in a telephony system. A subscriber A, called the far end subscriber below, is connected to a hybrid (a hybrid forms the interface between a four-wire and a two-wire connection, as is well known in the art) over a two-wire line. Similarly a subscriber B, called the near end subscriber below, is connected to another hybrid over a two-wire line. The two-wire lines transfer both incoming and outgoing speech signals. Outgoing speech from far end subscriber A is transferred to near end subscriber B over the upper two-wire line in Fig. 1. Similarly outgoing speech from near end subscriber B

25

is transferred to far end subscriber A on the lower two-wire line in Fig. 1. However, the lower two-wire line from subscriber B to subscriber A also contains an echo of outgoing speech from subscriber A, which the hybrid at subscriber B was not able to suppress completely. Similarly the upper two-wire line in Fig. 1 contains echo from outgoing speech from subscriber B.

Fig. 2 illustrates how the echo back to subscriber A is cancelled at the near end side (a similar arrangement is provided at the far end side). Input signal $x(n)$, where n denotes discrete time, represents speech from subscriber A. The input signal $x(n)$ is attenuated by the hybrid, represented by a filter 10 with transfer function $H(q^{-1})$ and a summation unit 14, and the resulting echo signal $s(n)$ is combined with the near end signal $v(n)$, which may or may not contain near end speech, in summation unit 14. The attenuation of filter 10 is represented by the echo path attenuation ERL (ERL = Echo Return Loss). Thus, the resulting output signal $y(n)$ contains both the near end signal and echo from the far end signal. Furthermore, input signal $x(n)$ is also forwarded to an adaptive filter 12, which models the impulse response of the hybrid by adjusting its filter coefficients. The resulting estimate of echo signal $s(n)$ is denoted $\hat{s}(n)$. This estimate is, in a summation unit 16, subtracted from output signal $y(n)$ (ERLE = Echo Return Loss Enhancement represents the obtained improvement in echo attenuation), and the resulting error signal $e(n)$ is forwarded to adaptive filter 12 for adjustment of the filter coefficients and to the two-wire line back to far end subscriber A.

A problem with the simple block diagram of Fig. 2 is that signal $y(n)$ may contain, in addition to the echo signal $s(n)$, a speech signal $v(n)$ from subscriber B. This situation is called double-talk. During double-talk adaptive filter 12 will try to model not only the echo signal $s(n)$ but also the speech signal $v(n)$. Thus, the adaption of filter 12 must be controlled during double-talk.

Figure 3 illustrates a block diagram of a dual filter echo

canceller described in [1, 2] intended to solve this double-talk problem. Adaptive filter 12 is continuously updated whether there is double-talk or not. However, in this case the output from summation unit 16 is only forwarded to adaptive filter 12 and not to the two-wire line back to far end subscriber A. Instead the actual echo cancellation is performed by a programmable foreground filter 18, which forwards an echo estimate to a summation unit 22, which forwards a resulting error signal $e_e(n)$ to the two-wire line back to far end subscriber A. The coefficients from the adaptive background filter 12 are transferred to the programmable foreground filter 18 whenever the adaptive background filter 12 is considered better than the programmable foreground filter 18. This usually occurs when there is no double-talk. During double-talk the coefficients that were transferred to the programmable foreground filter 18 just before the double-talk situation occurred are kept for echo cancellation during the double-talk period. Once the double-talk situation no longer exists and the adaptive background filter 12 is determined to give better performance, filter coefficients are once again transferred from filter 12 to filter 18.

The method to compare the performance of the two filters described in [1, 2] may be summarized as follows. The main idea is to compare the residual energy from the two filters. Thus, filter coefficients are transferred only if

$$E|e_b(n)| < \mu \cdot E|e_f(n)| \quad (1)$$

where $E(.)$ denotes estimated residual energy level and μ is a constant, which is chosen to 7/8 in [1]. In order to make this algorithm perform well the following two requirements are necessary

$$E|e_b(n)| < \lambda \cdot E|y(n)| \quad (2)$$

$$E|y(n)| < E|x(n)| \quad (3)$$

where λ is a constant, which in [1] equals 1/8 (corresponding to -18 dB). If the above three conditions are fulfilled the filter

coefficients of filter 12 are transferred to filter 18.

Equation (1) above means that the residual echo energy level from the background filter 12 should be lower (by a factor μ) than the residual energy from the foreground filter 18. Condition (2) means that the echo return loss enhancement (ERLE) must have reached a predetermined threshold of $-20 \log \lambda$ dB. Condition (3) means that there should not be an obvious double-talk situation (if $y(n)$ has more energy than $x(n)$ it must contain something in addition to the echo signal $s(n)$, namely near end speech). As a further condition it may be required that the above three conditions are simultaneously fulfilled for a predetermined time period, for example 48 ms.

Since the configuration of [1, 2] only uses the programmable foreground filter 18 for actual echo cancellation, it is very important that the adaptive filter 12 is always transferred when it performs better. However, due to the problems stated below this may not always occur.

One problem occurs if the near end side has a high background noise level. In this case the residual echo $e_r(n)$ will be buried in noise. This means that condition (1) above becomes blind; no incentive is given to transfer the background filter to the foreground filter.

Another problem is that condition (2) requires that the echo return loss enhancement ERLE should have reached 18 dB before any transfer of the background filter can take place. However, this situation may never be achieved if the background noise level is high and the echo return loss (ERL) is also high.

A further problem is that the ERLE requirement of 18 dB may never be fulfilled if the echo path has a high degree of non-linearity.

Since the adaptive filter 12 of [1, 2] is allowed to adapt continuously it will diverge from its optimum during double-talk.

This divergence is not restored, which means that the adaptive filter needs a new convergence period after every double-talk situation before it reaches the same performance as the programmable filter. This implies that the convergence process of the echo canceller will become very inefficient in a fast alternating duplex situation.

Figure 4 illustrates an echo canceller using the method of the present invention. In the echo canceller of Fig. 4 filter 12 is an adaptive filter and filter 18 is a programmable filter, as in the prior art echo canceller of Fig. 3. However, in the echo canceller of Fig. 4 the two filters are used completely in parallel, i.e. residual signals $e_a(n)$ and $e_p(n)$ are obtained for both filters, and a decision logic 24 decides which signal to choose as the actual output signal $e(n)$. Furthermore, as indicated by double arrow 21, both filters may be transferred or copied.

In accordance with the present invention decision logic 24 uses the quality measure

$$q_i = \frac{E\hat{s}_i(n)y(n)}{Ee_i^2(n)} \quad (4)$$

where $i=a,p$, to decide which residue signal $e_a(n)$ or $e_p(n)$ to use as the actual output signal. This choice of quality measure will now be explained.

Consider the signal

$$y(n) = s(n) + v(n) \quad (5)$$

where $s(n)$ represents the echo signal and $v(n)$ represents near-end noise and speech. From (5) it can be seen that the numerator of (4) is a correlation between the estimated echo and the true echo, with near-end speech and noise added. This correlation will be high if the filter is well adjusted to the echo path. Since $\hat{s}_i(n)$ is independent of $v(n)$, the numerator of q_i will not vanish when the background noise level is high. However, since $Ee_i^2(n)$

is used as the denominator, q_i will decrease in the presence of near-end speech or noise. Thus, a convenient condition for decision logic 24 to select residual signal $e_a(n)$ as the "best" signal, is to require that

$$q_a > Aq_p + B \quad (6)$$

5 is fulfilled. Here A is a predetermined factor and B is a predetermined offset.

To avoid selecting the adaptive filter during an obvious double-talk situation, it may also be required that the following condition is fulfilled

$$q_a > C \text{ OR } [E y^2(n) < \alpha \cdot NL \text{ AND } q_a > B] \quad (7)$$

10 before the adaptive filter is selected as the best filter. Here C represents an offset which is greater than offset B. Furthermore, α is a factor and NL is the measured noise level.

Figure 5 illustrates an embodiment of the method in accordance with the present invention in which the quality measure '(4) is
15 used to determine the best filter. In step 500 the next sample is used to calculate new quality measures in steps 510 and 520. Step 530 performs the test in accordance with condition (6). If condition (6) is fulfilled, step 540 tests the first part of condition (7). If this test fails the alternative branch 550
20 including the second part of condition (7) is tested. If either of tests 540, 550 is successful the algorithm proceeds to step 560. This step tests whether the following condition is fulfilled

$$E e_p^2(n) < \beta \cdot E e_a^2(n) \quad (8)$$

where β is a predetermined factor. This step tests whether the
25 programmable filter has a lower residual signal energy than the adaptive filter. If this is not the case the adaptive filter is selected as the output filter, and this filter is used to produce the actual output signal $e(n)$. On the other hand, if test 560 indicates that the programmable filter actually has a smaller

residual signal energy, this filter will be used to produce the output signal in step 580. Similarly the programmable filter will be used if the test in step 530 fails and if both tests 540 and 550 fail.

5 In a preferred embodiment of the method illustrated in Fig. 5, the following values have been used for the various predetermined constants.

A=2

B=0

10 C=1

$\alpha=10$

$\beta=1$

15 With these values it may be seen that condition (6) is less conservative than the conditions in [1, 2]. For example, C=1 implies that in the stationary case ERLE should be higher than 0 dB. This is much lower than the value 18 dB in [1, 2]. This condition is further relaxed to $q_n > 0$ when $Ey^2(n)$ falls below the noise level.

20 Figure 6 illustrates a preferred embodiment of the method in accordance with the present invention. In this preferred embodiment steps 500-560 are the same as in the embodiment of Fig. 5. However, instead of using the selected filter directly to produce an output signal, this preferred embodiment makes a smooth transition from one filter to the other by linearly
25 combining the residual signals from the two filters in accordance with step 620. Each time the adaptive filter is chosen as the best filter, a filter state variable FS is increased in accordance with step 600. Similarly, each time the programmable filter is selected as the best filter, filter state variable FS is
30 decreased in accordance with step 610. The calculated filter state variable FS is then used in step 620 to form a linear combination between residual signals $e_a(n)$ and $e_p(n)$. Here the variable τ represents a transition time, for example 128 sample periods. As can be seen from step 620 the proportion of a

selected filter will increase while the proportion of a non-selected filter will decrease. When a filter has been consistently selected for τ sample periods the smooth transition has been completed.

5 Step 620 performs a linear combination of $e_p(n)$ and $e_s(n)$. However, this is not absolutely necessary. For example, it is also possible to use non-linear weighting factors, although the linear combination is probably optimal.

10 A presently preferred embodiment of the method illustrated in Fig. 6 uses the same values for the predetermined constants A, B, C, α , β as the preferred embodiment of Fig. 5.

15 The methods illustrated in Figs. 4 and 5 are concerned with selecting and using the proper filter for producing the actual output signal $e(n)$. However, as indicated by the double arrow 21 in Fig. 4 each filter may also be transferred or copied to the other filter. For example, if the adaptive filter is consistently better than the programmable filter, it may be preferable to copy the coefficients of the adaptive filter to the programmable filter. On the other hand, after a double-talk situation, in
20 which the adaptive filter has diverged, it is probably a good idea to transfer the coefficients from the programmable filter to the adaptive filter, since the estimated echo of the programmable filter is probably better than the echo estimate of the diverged adaptive filter (the estimated echo before the double-talk situation is probably a good starting point for an adaption to a
25 new echo estimate after the double-talk situation).

30 Figure 7 illustrates a preferred embodiment of a method for transferring filter coefficients from one filter to the other which is based on the same algorithm as the filter selection methods of Figs. 5 and 6. Thus, steps 500-550 are the same as in Figs. 5 and 6. If the adaptive filter has been selected as the best filter a counter COUNT is incremented in step 700. Step 710 tests whether COUNT exceeds a predetermined constant T (for

example 2 047). If COUNT exceeds T, this means that the adaptive filter has been selected T times. Therefore the adaptive filter is copied to the programmable filter (step 730) and the counter COUNT is reset to zero (step 720). Thus, if the adaptable filter is consistently selected it will be transferred to the programmable filter.

On the other hand, if the programmable filter has been selected as the most appropriate filter, step 740 tests whether the following two conditions are both fulfilled

$$Ee_p^2(n) < \beta \cdot Ee_a^2(n) \text{ AND } Ee_a^2(n) > \gamma^2 \quad (9)$$

These conditions imply that the adaptive filter performs significantly worse (controlled by the factor β) than the programmable filter and that the residual energy must exceed a certain threshold γ^2 to avoid taking decisions on low non-significant energy levels. Suitable values are $\beta = \frac{1}{2}$ and $\gamma = -40$ dBm0. If step 740 is successful the programmable filter is copied to the adaptive filter (step 760) and counter COUNT is reset to zero (step 750).

The two situations described so far are the situations in which filter coefficients are actually copied. However, if test 710 fails the algorithm will proceed to step 790, which implies that no filter coefficients are copied. This occurs when the variable count has not yet reached the value T.

Another situation in which no filter coefficients are copied is when test 740 fails. In this situation the algorithm proceeds to step 770. Step 770 tests whether the following condition

$$Ey^2(n) > \alpha \cdot NL \quad (10)$$

is fulfilled. Thus, step 770 tests whether signal $y(n)$ exceeds the noise level. If this is the case there probably is a double-talk situation, since signal $y(n)$ probably contains speech and the adaptive filter does not perform significantly better than the programmable filter. Consequently the variable COUNT is reset

to zero in step 780 to indicate that this is certainly not the time to transfer the adaptive filter to the programmable filter. On the other hand, since step 740 failed, the programmable filter is not significantly better than the adaptive filter. Thus, none
5 of the filters is transferred (step 790).

Finally, if step 770 fails, this indicates that no decisions can be made, and things are left as they are (no filter is copied, COUNT is not changed).

10 In a preferred embodiment of the method illustrated in Fig. 7 the following constants are used:

A=1

B=0,125

C=1

$\alpha=10$

15 $\beta=\frac{1}{2}$

$\gamma^2=-40$ dBm0

It will be understood by those skilled in the art that various modifications and changes may be made to the present invention without departure from the spirit and scope thereof, which is
20 defined by the appended claims.

REFERENCES

[1] K. Ochiai et al, "Echo Canceller with Two Echo Path Models", IEEE Transactions on Communications, 25(6):589-594, June 1977

25 [2] US, A, 3 787 645

CLAIMS

1. A method of determining a quality measure representing performance of a filter in an adaptive echo canceller, characterized by the steps of:

5 estimating a correlation measure between an echo containing signal ($y(n)$) and an echo estimation signal ($\hat{s}(n)$) produced by said filter;

10 estimating a power measure of a residual signal ($e(n)$) formed by the difference between said echo estimation signal ($\hat{s}(n)$) and said echo containing signal ($y(n)$); and

 calculating said quality measure (q) by dividing said estimated correlation measure by said estimated power measure.

15 2. The method of claim 1, characterized in that said echo containing signal ($y(n)$) may contain, in addition to echo, noise and speech signals produced near said echo canceller.

3. An adaptive dual filter echo cancellation method in which an adaptive and a programmable filter (12, 18) are both used for estimating an echo signal, characterized by the steps of:

20 estimating a first correlation measure between an echo containing signal ($y(n)$) and an adaptive filter echo estimation signal ($\hat{s}_a(n)$);

25 estimating a first power measure of a first residual signal ($e_a(n)$) formed by the difference between said adaptive filter echo estimation signal ($\hat{s}_a(n)$) and said echo containing signal ($y(n)$);

 determining an adaptive filter quality measure (q_a) by dividing said estimated first correlation measure by said estimated first power measure;

30 estimating a second correlation measure between said echo containing signal ($y(n)$) and a programmable filter echo estimation signal ($\hat{s}_p(n)$);

 estimating a second power measure of a second residual signal ($e_p(n)$) formed by the difference between said programmable filter echo estimation signal ($\hat{s}_p(n)$) and said echo containing signal



(y(n));

determining a programmable filter quality measure (q_p) by dividing said estimated second correlation measure by said estimated second power measure;

5 comparing said adaptive filter quality measure (q_a) to said programmable filter quality measure (q_p) for determining whether said adaptive filter or said programmable filter gives the best estimate of said echo signal.

4. The method of claim 3, characterized by

10 selecting said adaptive filter as the filter that gives the best estimate of said echo signal only if the following condition (530) is fulfilled:

15 (i) said adaptive filter quality measure (q_a) exceeds the sum of a first predetermined offset (B) and the product of said programmable filter quality measure (q_p) and a predetermined first factor (A); and

selecting said programmable filter as the filter that gives the best estimate of said echo signal if the condition (i) is not fulfilled.

20 5. The method of claim 4, characterized by

selecting said adaptive filter as the filter that gives the best estimate of said echo signal only if at least one of the following further conditions (540, 550) is fulfilled:

5 (ii) said adaptive filter quality measure (q_a) is greater than a second predetermined offset (C), which is greater than said predetermined first offset (B), and

30 (iii) said adaptive filter quality measure (q_a) is greater than said first predetermined offset (B), and an estimated third power measure of said echo containing signal (y(n)) is less than the product of a measured noise level (NL) and a second predetermined factor (α); and

35 selecting said programmable filter as the filter that gives the best estimate of said echo signal if neither of conditions



(ii) and (iii) is fulfilled.

6. The method of claim 5, characterized by

selecting said adaptive filter as the filter that gives the best estimate of said echo signal only if the following further condition (560) is not fulfilled:

(iv) said estimated second power measure is less than the product of said estimated first power measure and a third predetermined factor (β); and

selecting said programmable filter as the filter that gives the best estimate of said echo signal if condition (iv) is fulfilled.

7. The method of claim 4, 5 or 6, characterized by using (570, 580) the selected filter for estimating said echo signal.

8. The method of claim 4, 5 or 6, characterized by combining (620) said first and second residual signals ($e_a(n)$, $e_p(n)$), increasing the proportion of the residual signal that corresponds to the selected filter and decreasing the proportion of the residual signal that corresponds to the non-selected filter.

9. The method of claim 7 or 8, characterized in that said first predetermined factor equals 2, said first predetermined offset equals 0, and said second predetermined offset equals 1.

10. The method of claim 5, characterized by

copying (760) said programmable filter to said adaptable filter if said programmable filter has been selected and the following conditions (740) are both fulfilled:

(iv) said estimated second power measure is less than the product of said estimated first power measure and a third predetermined factor (β),

(v) said estimated first power measure is greater than a predetermined constant (γ^2).

11. The method of claim 10, characterized by



WO 97/15124

PCT/SE96/01317

16

counting (COUNT) each time said adaptable filter has been selected; and

copying (730) said adaptive filter to said programmable filter when it has been selected a predetermined number (T) of times.

- 5 12. The method of claim 11, characterized in that said first predetermined factor equals 1, said first predetermined offset equals 0.125, and said second predetermined offset equals 1.

10