Self Adaptive Filtering of Noise from Speech: A Hearing Aid Application

D. Graupe , S.P. Basseas and J.K. Grosspietsch

Department of Electrical Engineering and Computer Science, University of Illinois at Chicago
and Intellitech Inc., Northbrook, IL 60062

Abstract

The paper discusses and evaluates a self adaptive filter of noise from speech. The filter employs a single microphone, and incorporates time-series identification algorithms with artificial intelligence decision procedures to discriminate between environmental noises, including babble from speech. The filter has recently been realized in a single CMOS microchip form which includes digital, analog and ADC circuits and is presently being used by seven hearing aid manufacturers in their otherwise conventional hearing aids, both of behind-ear and of inside-ear types. Whereas Communications applications of this adaptive filter have also been developed, the development of the hearing aid applications was emphasized noting that the hearing impaired have difficulty in understanding speech even without noise. Hence, this let alone in any kind of noise. Circuit and system design consideration are discussed and performance evaluation is presented, as is the relative performance as compared with conventional non-adaptive single microphone designs based on AGC and ASP-type compression circuits.

1. Introduction

The problem of understanding speech in noise affects any person, and even man-made speech recognizing systems. However, the hearing impaired who have difficulty in following speech even without noise, are far more susceptible to the effect of noise on understanding speech. The present paper discuss a self adaptive/self tuning filter of environmental noises, including babble (multi-speaker) noises and his colleagues from speech, which also possesses the capability of self adaptive speech/noise discrimination. This filter was developed by the first author for the general speech/noise problem but with a particular emphasis on the needs of the hearing impaired. Hence, requirements of single-microphone (the lady using her hearing aid cannot throw a second microphone on the passing truck or helicopter), of low supply voltage/low power (to facilitate using conventional button-type hearing aid batteries), and of small size to fit conventional in-the-ear or behind-the-ear hearing aids and even in-canal hearing aids.

There are many approaches to speech enhancement in the presence of noise. Most of them are nonadaptive and do not discriminate between speech and noise. The most common nonadaptive approach is that of reducing the gain of the hearing aid at frequencies below 500 or 700 HZ. This assumes that many environmental noises are of low frequency, which is often the case. However, this approach equally suppresses low-frequency speech cues and it does so regardless of the presence of noise. The hearing-impaired person, who typically has the greatest hearing loss in the high frequencies, is now also deprived of low-frequency cues. Meanwhile, high-frequency noise is unaffected.

Another nonadaptive approach is that of using compression circuits based on AGC (automatic-gain-control or automatic voltage control) and the related ASP compression circuits (where AGC is applied only to a range of frequencies). These systems suppress both noise and speech to the same degree since they have no speech/noise discrimination capabilities, nor are they able to identify and tune the speech or the noise parameters (see further discussion in section 3 below).

There are also several self-adaptive methods for the filtering of noise from speech. The most successful is based on Widrow's filtering theory (12), and on the rather similar Tsypkin theory (10). The Widrow-Tsypkin approach, while rigorous, is not totally adaptive in that it requires prior parameter knowledge of speech alone or of noise alone. This requirement limits the practical usefulness of the technique in hearing aid applications. It performance is good down to a SNR (signal-to-noise ratio) of -12 dB, yielding suppression of noise by up to 10 dB (2). Its performance, however, worsens considerably if the noise comes from the direction of the speech, as in a lecture hall or similar situation.

Among single-microphone adaptive filtering approaches is that of Weiss and Aschkenasy (11). This technique uses a cepstrum-like transformation and signal reconstruction. This method improves the SNR in white noise, but yields little intelligibility enhancement. This is partly due to poor reconstruction of unvoiced speech (6). It involves time delays for four FFT calculations (at least 512 points each), which constitute a major obstacle to hearing-aid applications.

Boll (1) reports similarly poor intelligibility enhancement for suppression of noise using direct spectrum subtraction of averaged speech spectra.

The method of Sambur (7), where the input signal of speech-plus-noise is delayed through a tapped delay line, is essentially a single-microphone extension of the Widrow-Tsypkin approach. It assumes that speech is highly correlated in contrast to noise. Thus, the correlated part of speech plus noise is considered to be speech, in order to provide the speech parameters for subsequent filtering. This philosophy is inadequate for unvoiced speech, which is stochastic in nature compared with voiced speech. It also fails when noise is correlated, as it often is. It is not surprising then that this method also does not improve intelligibility (2).

2. Description of the Self-Adaptive (ZETA) Filter

The present self adaptive filter (and its hearing aid version, the Zeta filter), which was first patented by Graupe and Causey in 1976, is the Zeta noise blocker (first patented in 1976) is a single-microphone self-adaptive filter whose block diagram is given in Figure 1. The input to the filter is y(t) where y(t)=s(t)+n(t): t=time, s(t) and n(t) denoting speech and noise. It operates in real time such that no delay exists between s(t) and the corresponding filtered speech output s(t).

The filter first identifies the parameters vector p(y) of the input y(t). Subsequently, parameter separation is performed, where the parameter vector p(y) is separated in sub-vectors of noise parameters and of speech parameters. The later parameter separation, as is justified theoretically in Section 11.2 of (4), is possible only when assuming that speech parameters have different features than noise parameters. Therefore, only signals whose parameter features are different from speech parameter features will be identified as noise and thus be filtered. This is the situation for most environmental noises, since their spectural or time-domain parameters in time relative to speech parameters. Indeed, speech phonemes have characteristic random-like changes, at intervals of 5 to 200 milliseconds, depending on their being voiced or unvoiced. Hardly any environmental noise displays such time variations. Even speech babble does not have the same characteristic time variation because of the acoustic averaging effect over many simultaneous conversations.

In the present realization, an artificial intelligence that considers features of speech parameters in speech recognition algorithms is employed.

Once parameter separation has been performed, a self-adaptive decision on the existence of noise n(t) in the input y(t) is possible. When no noise has been found to exist, the filter becomes transparent, such that the filter's output s(t) is set to be equal to y(t). Otherwise, the filter (which is essentially an adjustable Wiener filter) is automatically tuned to optimally filter out the noise that has been identified. Although the filter is a linear filter, its optimization differs from that of the Wiener filter. The difference is that it employs an artificial intelligence to heuristically modify the original Wiener filter's optimization by taking into account characteristics of speech phonemes: see Chapter 10 of (4). This artificial intelligence also facilitates heuristic adjustments for babble and white noise conditions, as identified by the filter (6).

The hearing-aid version (Zeta) of the adaptive filter (manufactured by Intellitech, Inc., Northbrook, IL) is presently realized in a 0.219 x 0.159 inch CMOS chip which includes an A/D on board. It is small enough to fit into any conventional BTE (behind-the-ear) hearing aid and a large percentage of ITE (in-the-ear) hearing aids. It uses standard hearing aid batteries (nominal 1.4 Vdc), including Zinc-Air and rechargeable batteries that can operate the circuit down to 1.1 Vdc, as is required for hearing aid applications. Its power consumption is below 1 mA. Thus it adds about one third to the current drain rate of the battery.

A more powerful version of the same filter for non-hearing-aid applications also has been constructed and tested by the present authors. This is the CAF (Communication Adaptive Filter). It employs the same principles and essentially same algorithms. However. it is not limited in its dynamic range to a 1.4 Vdc supply, nor is its size restricted to a single CMOS chip of the Zeta's present dimensions

3. Performance Evaluations

Clinical intelligibility tests on a prototype of the Zeta self-adaptive filter coupled with a Rexton 25 PP body hearing aid were performed in 1984 at the Siegal Ear Institute of the Michael Reese Medical Center, Chicago, IL (8). These tests were performed on 20 subjects, 5 with normal hearing, 5 with a mild-to-moderate sensorineural hearing loss, 5 with a moderate to severe sensorineural loss, and 5 with a steeply sloping high-frequency sensorineural loss.

Live voice readings of Northwestern University Test 6 (NU 6) were used with 5 noise conditions: 1) 600-800 Hz band-filtered noise (48 dB/octave); 2) 1700-2400 Hz band-filtered noise; 3) cafeteria noise; 4) six-speaker babbles, (3 male, 3 female); and, 5) white noise. Speech level was 68 dB SPL and noise level was varied by an up-down adaptive procedure to degrade listener's discrimination score to within the 30 to 50 percent range. Tests results are shown in Table 1.

We comment that a 70% intelligibility score with a randomly mixed set of single-syllable words as in Table 1, implies a much higher intelligibility score (around 100%) with conversational speech sentences. It should also be noted that the filtered scores should not exceed the score for the quiet condition since the filter cannot yield intelligibility scores (in the presence of noise) that are above those achieved for the same speech without any noise. For example, a 69.6 percent score in column "F" of Table 1 for Group 3 with cafeteria noise implies a restoration of 89.7% of the maximum ("in quiet") score. For this group and noise condition without the filter (column NF), the average score was 36.0 or 46.4 percent of the maximum ("in quiet") score. The filter thus facilitated regaining of 89.7% - 46.4% = 43.3% of the maximum ("in quiet") score. The maximum possible score is 100 percent and the maximum attainable improvement is 100 - 46.4 = 53.6%. Hence, the percentage of maximum attainable score improvement provided by the filter was 43.3/53.6 namely, 80.8 percent under this noise condition. The results of the various noises for the 4 groups in terms of improvement as a percentage of maximal-attainable improvement are shown in Table 2.

Results of tests on the Zeta self-adaptive filter in actual chip-form when factory assembled in a commercial BTE hearing aid are reported by Wollinsky (13) and by Ellison (3). Wollinsky's results are reproduced in Table 3. These results are for 18 hearing-impaired patients all in the moderate/severe loss group, using the same Northwestern University Monosyllabic (NU6) test tapes as for the data of Tables 1 and 2. The results by Edison (3) are given in Table 4.

When comparing the scores of Tables 3 or 4 with those of Table 2, we note the Zeta noise-blocker chip used in the 1986 tests (13) was an improved version of the prototype used in the 1984 tests (8). As a result, slightly better performance was obtained in the 1986 study. Note also that the later study (13) used a BTE hearing aid while the earlier study (8) used a body aid. In both tests, improvements for low-frequency (600-800 Hz) noises were higher than for high-frequency noises 1700-2400 or 2700-3500 Hz. This is to be expected because high-frequency noises are less likely to mask speech than low-frequency noises.

Test results performed by R. LaRose of Hinsdale, Illinois on the same prototype of the filter used by Stein and Dempsey-Hart (8) are given in Table 5. Intelligibility scores at various SNR ranging from -20 to +5 DB are shown. These data indicate that the filter performs well on normal-hearing persons at - 20 dB SNR, in contrast to the methods discussed in the Introduction and Background section of this paper.

The extended (CAF) version of the Zeta self-adaptive filter is compared in terms of intelligibility scores and of SNR improvement with the single-chip version of the filter. The data are summarized in Tables 6 and 7, respectively.

These data indicate an average improvement in performance of the extended filter over the hearing aid version by 35 to 40 percent.

The filter's performance is illustrated graphically in Figure 2 in terms of the signal's time behavior. The top trace shown is the speech signal s(t), the second trace shows the noise n(t), the third trace shows speech and noise y(t)=s(t)+n(t), and the lowest trace sows the output of the filter s(t). The traces are for exactly the same data and filtering run.

Observe that compression-type non-adaptive filters, by their non-adaptive nature and by their lack of ability to discriminate between noise and speech, do not provide effective filtering when noise is present, whereas, they confuse speech with noise when speech is at above some threshold level. This is illustrated in Fig. 3 below for tests with the same speech and noise conditions, as performed on hearing aid, using the Zeta version of the present filter as compared with hearing aids using compression-type non-adaptive filtering, denoted as AGC and as ASP (limited frequency-band AGC) in the notation of Fig. 3 We comment that ideally all filters should aim at yielding an output as close as possible to the top trace of Fig. 2(a). Note the lack of transparency of non-adaptive methods vs. adaptive (Zeta) filter under no-noise conditions in Fig. 2(a) where compression-based methods do confuse noise with speech, and lack of improvement (actually, often deterioration) in signal to noise ratio (SNR) for non-adaptive filters relative to the input signal ("Unfiltered" top trace of Fig. 2(b), (c), (d) below), and in comparison to the significant SNR improvement in all cases for the adaptive filter ("Zeta" - bottom trace in Fig. 2(b), (c), (d) below).

4. Conclusions

The self-adaptive noise filter described above, in its hearing aid (Zeta) version and in its communications (CAF) version, incorporate self adaptive filtering and parameters identification algorithms combined with artificial-intelligence-based discrimination algorithms to result in a real-time single-microphone self adaptive filter. It has been realized with mixed digital and analog circuitry in a single CMOS (Zeta) chip that fits both conventional inside-ear and behind-ear hearing aid using supply voltages from conventional hearing aid batteries down to IIVDC at below mA. Presently, seven hearing aid manufacturers incorporate the Zeta version of this filter in their commercially-available in-ear and behind-ear hearing aids.

The resulting adaptive filter shows significant improvements in speech intelligibility and in SNR over a wide range of environmental noises, including babble (multi-speaker) noise, (see Tables 1 to 7, describing several independent real-time/real-patient tests), and over a wide range of input (infiltered) SNR (as low as -20 dB, see Table 5). The performance was tested with both hearing impaired and normally-hearing persons using in most tests actual hearing aids.

The results show and explain also the vast difference in performance between the adaptive filter and non-adaptive compression type filters, known as AGC and ASP (compression) filters, as shown in Fig. 2. The latter results illustrate the non-transparency of non-adaptive filters when no noise is present vs. the transparency of the adaptive filter. They also illustrate the lack of improvement (in fact, deterioration) in SNR for non-adaptive filters in the presence of noise, against the improvements obtained with the adaptive filter under exactly the same conditions.

References

1. Boll, S.F., Suppression of acoustic noise in speech using spectral subtraction. IEEE Trans Acoust Speech & Signal Proc. ASSP-29: 113-120, 1979.

2. Christiansen, R.W., Chabries, D.M., and Anderson, D., Noise reduction in speech using a modified LMS adaptive predictive filter. Proc. IEEE 7th Conf. of the Eng. in Med. & Biol. Sco., 1985.

3. Ellison, C., Proceedings of the Michigan Speech and Hearing Assoc. Meeting, March 21, 1987.

4. Graupe, D., Time Series Analysis, Identification and Adaptive Filtering, Malabar, FL: Krieger Publishing Co., 1984.

5. Graupe, D., Grosspietsch, J.K., and Basseas, S.P., A self adaptive filter of environmental noises from speech and its evaluation. Proc. IEEE CDC Conf., 1985.

6. Graupe, D., Grosspietsch, J.K., and Basseas, S.P., Self-adaptive filtering of environmental noises from speech. proc. AIAA/IEEE 6th Avionics Sys. Conf., Baltimore, MD, 1984.

7. Sambur, M.R., Adaptive noise cancelling for speech signals. IEEE Trans Acoust Speech & Sign Proc ASSP-26:419-423, 1979.

8. Stein, L. and Dempsey-Hart, D., Listener-assessed intelligibility of a hearing aid self-adaptive noise filter. Ear and hearing 5(4):199-204, 1984.

9. Tsypkin, Y.Z., Foundations of the Theory of Learning, Academic Press, NY, 1973.

10. Weiss, M.R. and Aschkenasy, E., Wideband speech enhancement. RADC-TR-81-53, Final Tech Rep., May 1981.

11. Widrow, B., et al., Adaptive noise cancelling principles and applications. Proc IEEE 1692-1716, Dec. 1975.

12. Wollinsky, S., Clinical assessment of a self adaptive noise filtering system. The Hearing Journal 29-32, Oct. 1986.

Table 2.
Intelligibility improvement as a percent of maximum attainable improvement (11). The intelligibility improvement is equal to 100 · (F-NF)/(Quiet-NF), where the NF, F and Quiet conditions are as shown in **Table 1**.

	600/ 800Hz	1700/ 2400Hz	Cafeteria	Babble	White Noise
Group 1	55.1	31.9	50.7	31.6	8.3
Group 2	75.7	28.4	55.6	45.0	22.8
Group 3	66.0	21.3	80.8	21.0	−2.1
Group 4	36.5	38.8	12.5	32.1	34.3

Table 1.
Intelligibility Tests Data for Four Groups of Subjects (8). Monosyllabic word intelligibility is shown averaged over 5 subjects in each group.

		600/800 Hz		1700/2400 Hz		Cafeteria		Babble		White Noise	
	In Quiet	NF	F	NF	F	NF	F	NF	F	NF	F
Group 1 Normals	97.6	44.2	73.6	40.0	58.4	38.4	68.4	36.8	56.0	39.2	44.0
Group 2 mild/ moderate	94.4	36.4	80.4	38.0	54.0	32.6	69.2	38.4	63.6	40.0	52.4
Group 3 moderate/ severe	77.6	37.6	64.0	40.0	48.0	36.0	69.6	37.6	46.0	39.2	38.4
Group 4 sloping high-freq. loss	74.8	40.8	53.2	40.8	54.0	42.8	46.8	42.4	52.8	34.0	48.0

Key: NF-nonfiltered, F-filtered

Table 3.
Intelligibility-improvement as a percent of maximum attainable improvement. Data are for a BTE hearing aid (13).

	Ave. Score	Percent Intelligibility Improvement			
18 patient-average	In Quiet	600/ 800Hz	2700/ 3500Hz	Cafeteria	Babble
Moderate/ Severe loss	66.1	82.0	43.0	55.0	23.0

Table 4.: MEAN SPEECH DISCRIMINATION SCORES (%) IN PRESENCE OF BACKGROUND NOISE (Ref. 3)

NOISE TYPE	FILTER OFF	FILTER ON
Low Frequency (600-800 Hz)	35	63
Cafeteria	33	60
Babble	33	57
High Frequency (2700-3500 Hz)	50	60

Table 5. Intelligibility Scores at Various SNR.

Patient	Speech dB (HL)	Noise dB (HL)	SNR dB	Noise Type Hz	Score Non-Filtered	Filtered	Word List
Mrs. N.	65	60	+5	600/800	16 %	60 %	W-22
Normal	50	70	−20	600/800	24 %	80 %	NU-6
Mr. R.	60	65	−5	600/800	4 %	32 %	NU-6
Mr. N.	65	65	0	Babble	28 %	60 %	NU-6
Mrs. W.*	67.5	62.5	−5	Babble	0 %	24.5%	CID Sentences
Normal	50	70	−20	Babble	20 %	52 %	NU-6
Mr. N.	60	60	0	Cafeteria	36 %	84 %	W-22
Mrs. O.	57	67	−10	Cafeteria	0 %	64 %	W-22
Mrs. O.	57	62	−5	Cafeteria	64 %	94 %	CID Sentences
Mrs. O.	57	67	−10	Cafeteria	0 %	64 %	W-22
Mr. N.	60	65	−5	White Noise	0 %	16 %	W-22
Mrs. W.*	65	62.5	+2.5	White Noise	5 %	66 %	CID Sentences
Mrs. W.*	50	60	−10	2.4/3.0 KHz	45 %	92 %	CID Sentences

Mr. N.—Moderate to severe hearing loss 30–65 dB
Mr. R.—Mild loss 15–25 db, discrimination difficulties
Mrs. O.—Severe loss 55–75 db, low discrimination
*Mrs. W.—Mrs. W. scored nearly 0 at no noise with word lists, hence tested with sentences

Table 6. Intelligibility test scores: NU-6 monosyllabic word list.

Type of Noise	CAF 2 Unfil*	Fil**	Zeta (1) Unfil	Fil
250–400 Hz	28%	72%	44.2%	73.6%(2)
1200–1700 Hz	36%	48%	52 %	64 %(3)
	16%	36%	72 %	72 %(4)
2400–3000 Hz	68%	74%	92 %	100 %(4)
Babble	26%	44%	36.8%	56 %
Cafeteria	32%	90%	38.4%	68.4%

(% of fully correct word recognition)
*Unfiltered **Filtered
(1) For people with normal hearing, results from [11]
(2) Test for 600–800 Hz noise
(3) Tests by R. LaRose Noise of 1200–1400 Hz
(4) Test by R. LaRose

Table 7.

Improvement in signal-to-noise ratio (in dB)

Filter type	Improvement via Filter in dB CAF		Zeta	
Type of Noise	Low Noise	High Noise	Low Noise	High Noise
250–400 Hz	+20.6	+25.5	+15.1	
2400–3000 Hz	+25.5		+10.35	
Babble	+ 8.4	+ 6.7	+ 6.3	+ 2.6
Cafeteria	+12.2	+16.2	+11.6	+12.7

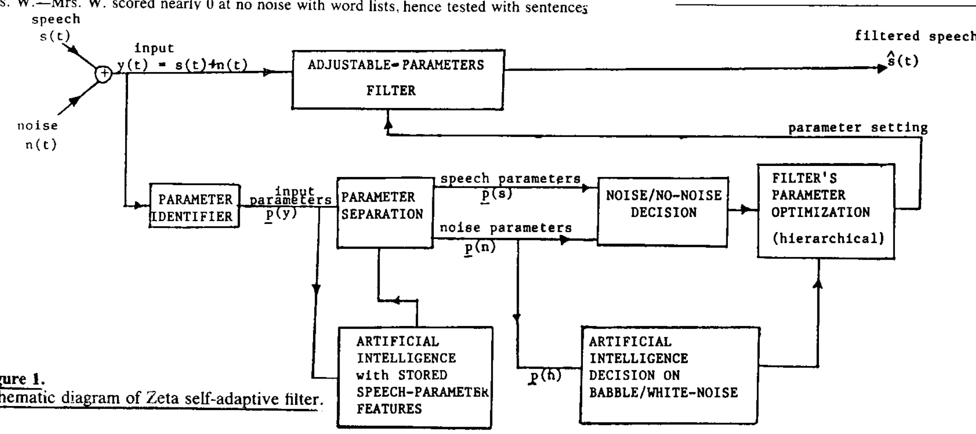


Figure 1.
Schematic diagram of Zeta self-adaptive filter.

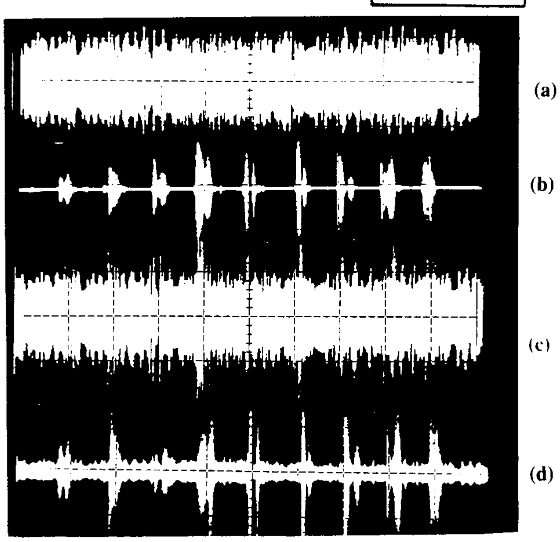


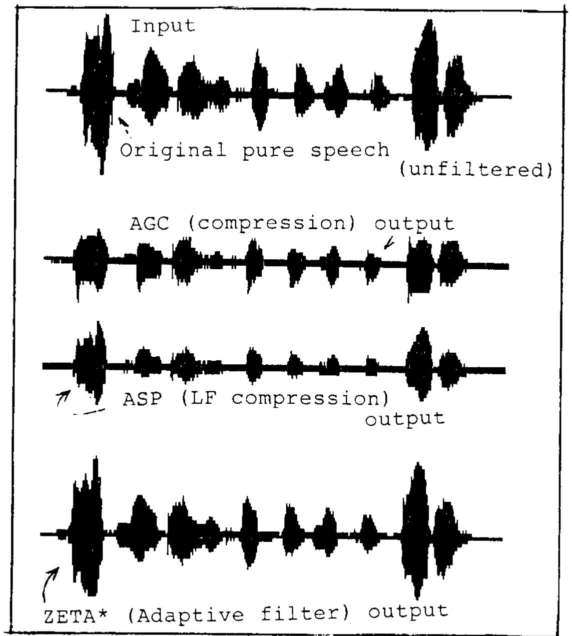
Figure 2.
Scope traces of filter's performance: Trace (a) shows 400–600 Hz Noise = n(t) of Fig. 1

Trace (b) shows speech (counting from 2 to 10) = s(t) of Fig. 1.

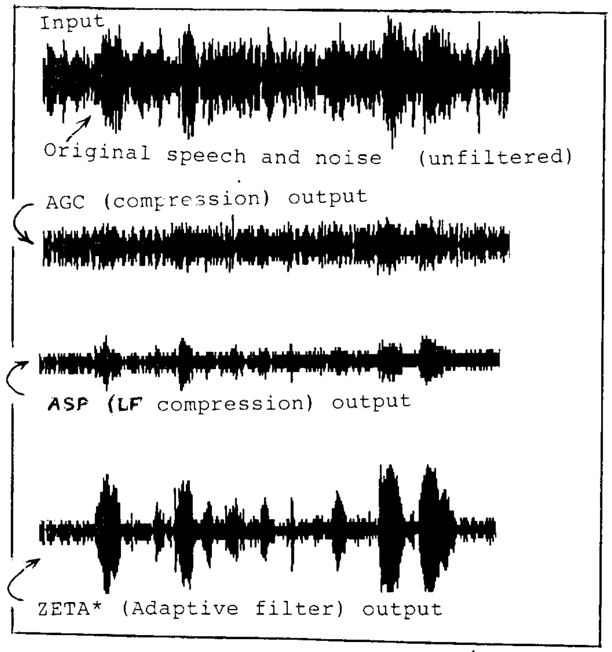
Trace (c) shows speech plus noise at filter's input = y(t) of Fig. 1.

Trace (d) shows speech filtered by the Zeta self-adaptive filter = ŝ(t) of Fig. 1

All traces were recorded simultaneously in real time.

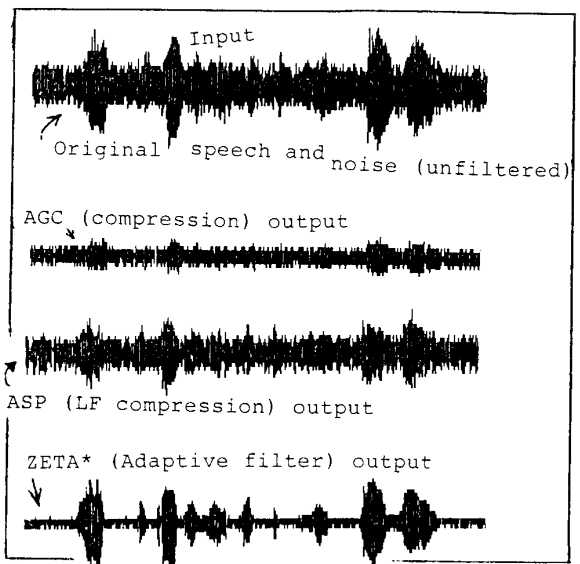


(a) Speech without any Noise

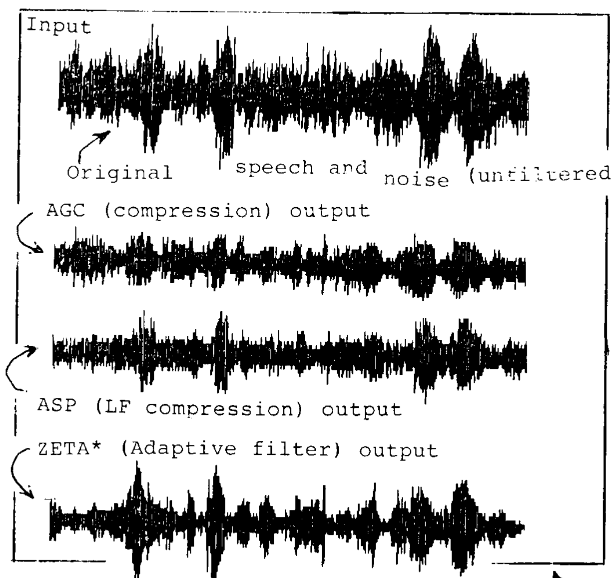


(b) Speech in Low Freq. Noise

*ZETA is the trade mark of Intellitech Inc, 900 Skokie Blvd.,Northbrook,IL. 60062



(c) Speech in High Freq. Noise



(d) Speech in Babble(4 speaker) Noise

Figure 3.
Comparison of Adaptive Filtering and Non Adaptive (compression-type) Filtering Performance for Various Noises.
Top trace of (a),(b),(c),(d) is the exact input of the 3 filters considered, whose outputs are the 3 lower traces of (a),(b),(c) and (d) respectivly.