We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Detection of echo generated in mobile phones

Tõnu Trump
Ericsson AB
Sweden

1. Introduction

Echo is a phenomenon where part of the sound energy transmitted to a receiver reflects back to the sender. In telephony it usually happens because of acoustic coupling between the receiver's loudspeaker and microphone or because of reflections of signals at the impedance mismatches in the analogue parts of the telephony system. In mobile phones one has to deal with acoustic echoes i.e. the signal played in the phones loudspeaker can be picked up by microphone of the same mobile phone.

People are used to the echoes that surround us in everyday life due to e.g. reflections of our speech from the walls of rooms where we are located. Those echoes arrive with a relatively short delay (in the order of milliseconds) and are, as a rule, attenuated. In a modern telephone system on the other hand the echoes may return with a delay that is not natural for human beings. The main reason for delay is in those systems signal processing like speech coding and interleaving. For example in a PSTN to GSM telephone call the one way transmission delay is around 100 ms making the echo to return after 200 ms. Echo that returns with this long delay is very unnatural to a human being and makes talking very difficult. Therefore the echo needs to be removed.

Ideally the mobile terminals should handle their own echoes in such a way that no echo is transmitted back to the telephony system. Even though many of the mobile phones currently in use are able to handle their echoes properly, there are still models that do not. ITU-T has recognized this problem and has recently consented the Recommendation G.160, "Voice Enhancement Devices" that addresses these issues (ITU-T G.160). Following this standard we concentrate on the scenario where the mobile echo control device is located in the telephone system.

It should be noted that differently from the conventional network- or acoustic echo problem (Sondhi & Berkley 1980; Signal Processing June 2006), where one normally assumes that the echo is present, it is not given that any echo is returned from the mobile phone at all. Therefore, the first step of a mobile echo removal algorithm should be detection of the presence of the echo, as argued in (Perry 2007). A simple level based echo detector is also proposed in (Perry 2007).

To design a mobile echo detector we first examine briefly the Adaptive Multi Rate (AMR) codec (3GPP TS 26.090) in Section 2. In Section 3 we present our derivation of the detector, which is followed by its performance analysis in Section 4. Some practicalities are explained in Section 5. Section 6 summarizes our simulation study.

Following the terminology common in mobile telephony, we use the term downlink to denote the transmission direction toward the mobile phone and the term uplink for the direction toward the telephony system.

2. Problem formulation

In order to detect the echo, which is a (modified) reflection of the original signal one needs a similarity measure between the downlink and the uplink signals. The echo path for the echo, generated by the mobile handsets is nonlinear and non-stationary due to the speech codecs and radio transmission in the echo path, which makes it difficult to use traditional linear methods like adaptive filters, applied directly to the waveform of the signals. As argued in (Perry 2007), the proper echo removal mechanism in this situation is a nonlinear processor, similar to the one that is used after the linear echo cancellation in ordinary network echo cancellers. In addition, as our measurements with various commercially available mobile telephones show, a large part of popular phone models are equipped with proper means of echo cancellation and do not produce any echo at all. Invoking a nonlinear processor based echo removal in such calls can only harm the voice quality and should therefore be avoided. That's why the first step of any mobile echo reduction system that is placed in the telephone system should be detection of the presence of echo. The nonlinear processor should then be applied only if the presence of echo has first been established.

Another important point is that speech traverses in the mobile system in coded form and that's why it is advantageous, if our detector were able to work directly with coded speech signals. Herein we therefore attempt to design a detector that uses the parameters present in coded speech to detect the presence of echo and estimate its delay. Exact value of the delay associated with the mobile echo is usually unknown and therefore needs to be estimated. The total echo delay builds up of the delays of speech codecs, interleaving in radio interface and other signal processing equipment that appear in the echo path together with unknown transport delays and is typically in the order of couple of hundreds of milliseconds.

The problem addressed herein is that the simple level based echo detector is not always reliable enough due to the impact of signals other than echo. The signals that are disturbing for echo detection originate from the microphone of the mobile phone and are actually the ones telephone system is supposed to carry to the other party of the telephone conversation. This is usually referred to as double talk problem in the echo cancellation literature. In this chapter we propose a detector that is not sensitive to double talk as shown in sequel of the chapter.

Let us now examine the structure of the AMR speech codec that is the codec used in GSM and UMTS mobile networks. The AMR codec switches between eight modes with different bit-rates ranging from 4.75 kbit/s to 12.2 kbit/s to code the speech signal. According to (3GPP TS 26.090), the AMR codec uses the following parameters to represent speech. The

Line Spectrum Pair (LSP) vectors, which are transformation of the linear prediction filter coefficients that have better quantization properties. The fractional pitch lags that represent the fundamental frequency of speech signal. The innovative codevectors that are used to code the excitation signal. And finally there are the pitch and innovative gains. In the detector, the LSP vectors are converted to the Linear Prediction (LP) filter coefficients and interpolated to obtain LP filters at each subframe. Then, at each 40-sample subframe the excitation is constructed by adding the adaptive and innovative codevectors scaled by their respective gains and the speech is reconstructed by filtering the excitation through the LP synthesis filter. Finally, the reconstructed speech signal is passed through an adaptive postfilter.

The basic structure of the decoder in a simplified form but sufficient for our purposes is shown in Figure 1 and described by the equation (1).

$$c \times g_c \times \frac{1}{1 + g_p z^{-T}} \times \frac{1}{A(z)} \times \frac{A(z/\gamma_n)}{A(z/\gamma_d)} \tag{1}$$

In the above c denotes the innovative codevector, g_c denotes the innovative gain (fixed codebook gain), g_p is the pitch gain, γ_n and γ_d are the postfilter constants and $A(z)$ denotes the LP synthesis filter. T is the fractional pitch lag, commonly referred to as "pitch period" throughout this chapter.
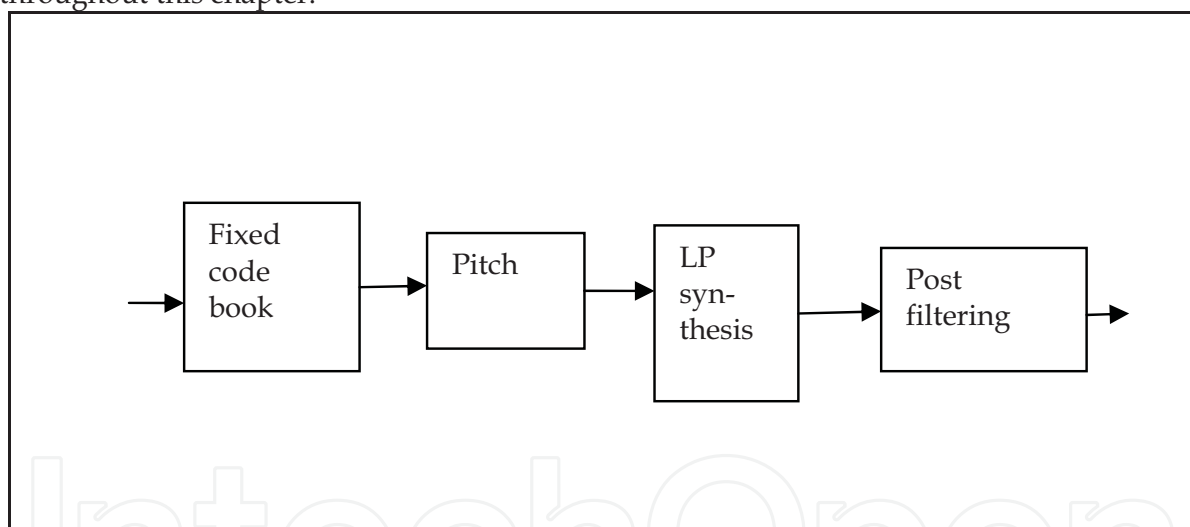

Fig. 1. Simplified structure of AMR decoder

Of the parameters present in AMR coded bit-stream, the pitch period or the fundamental frequency of the speech signal is believed to have the best chance to pass a nonlinear echo path unaltered or with a little modification. An intuitive reason for this is that a nonlinear system would likely generate harmonics but it would not alter the fundamental frequency of a sine wave passing it. We therefore select the pitch period as the parameter of interest.

3. Derivation

In this section we derive a structure for the echo detector based on comparison of uplink and downlink pitch periods. The derivation follows the principles of statistical hypothesis testing theory described e.g. in (Van Trees 1971; Kay 1998).

Denote the uplink pitch period for the frame t as $T_{ul}(t)$ and the downlink pitch period for the frame $t - \Delta$ as $T_{dl}(t - \Delta)$. The uplink pitch period will be treated as a random variable due to the presence of pitch estimation errors and the contributions from the true signal from mobile side.

Let us also denote the difference between uplink and downlink pitch periods as

$$w(t, \Delta) = T_{ul}(t) - T_{dl}(t - \Delta) \tag{2}$$

Then we have the following two hypotheses:

H_0: the echo is not present and the uplink pitch period is formed based only on the signals present at the mobile side

H_1: the uplink signal contains echo as indicated by the similarity of uplink and downlink pitch periods

Under hypothesis H_1, the process, w, models the errors of echo pitch estimation made by the speech codec residing in mobile phone but also the contribution from signal entering the microphone of the mobile phone. Our belief is that the distribution of the estimation errors can be well approximated by the Laplace distribution and that the contribution from the microphone signal gives a uniform floor to the distribution function. Some motivation for selecting this particular model can be found in Section 6.1.

We thus assume that under the hypothesis H_1 the distribution function of w is given by

$$p(w|H_1) = \begin{cases} \alpha \max\left(\dfrac{1}{2\sigma}\exp\left(-\dfrac{|T_{ul}(t) - T_{dl}(t - \Delta)|}{\sigma} \right), \dfrac{\beta}{b-a} \right), & a < w < b \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

The constant β, in the above equation, is a design parameter that can be used to weight the Laplace and uniform components and σ is the parameter of Laplace distribution. The variables a and b are determined by the limits in which the pitch period can be represented in the AMR codec. In the 12.2 kbit/s mode the pitch period ranges from 18 to 143 and in the other modes from 20 to 143. This gives us limits for the difference between uplink and downlink pitch periods a = -125 and b = 125 in 12.2 kbit/s mode and a = -123, b = 123 in all the other modes. α is a constant normalizing the probability density function so that it integrates to unity. Solving

$$\int_a^b p(w)\,dw = 1 \tag{4}$$

for α we obtain

$$\alpha = \frac{b - a}{2\sigma\beta\left(\ln\dfrac{2\sigma\beta}{b-a} - 1 \right) + (1 + \beta)(b - a)}. \tag{5}$$

Equation (3) can be rewritten in a more convenient form for further derivation

$$p(w|H_1) = \begin{cases} \dfrac{\alpha}{2\sigma}\exp\left(-\dfrac{\min\left(\left|T_{ul}(t)-T_{dl}(t-\Delta)\right|,-\sigma\ln\dfrac{2\delta\beta}{b-a}\right)}{\sigma}\right), & a < w < b \\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Under the hypothesis H_0, the distribution of w is assumed to be uniform within the interval $[a, b]$,

$$p(w|H_0) = \begin{cases} \dfrac{1}{b-a}, & a < w < b \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

We assume that the values taken by the random processes $w(t)$ at various time instances are statistically independent. Then the joint probability density is product of the individual densities

$$\begin{cases} p(\mathbf{w}|H_1) = \displaystyle\prod_{t=1}^{N} p(w(t)|H_1) \\ p(\mathbf{w}|H_0) = \displaystyle\prod_{t=1}^{N} p(w(t)|H_0). \end{cases} \tag{8}$$

Let us now design a likelihood ratio test (Van Trees 1971) for the hypotheses mentioned above. We assume that the cost for a correct decision is zero and the cost for any fault is one. We also assume that both hypotheses have equal a priori probabilities. Then the test is given by

$$\Lambda(T_{ul}) = \frac{\displaystyle\prod_{t=1}^{N}\dfrac{\alpha}{2\sigma}\exp\left(-\dfrac{\min\left(\left|T_{ul}(t)-T_{dl}(t-\Delta)\right|,-\sigma\ln\dfrac{2\delta\beta}{b-a}\right)}{\sigma}\right)}{\displaystyle\prod_{t=1}^{N}\dfrac{1}{b-a}} \underset{H_0}{\overset{H_1}{\gtrless}} 1, \tag{9}$$

Taking the logarithm and simplifying the above we obtain the following test

$$\sum_{t=1}^{N} -\min\left(\left|T_{ul}(t)-T_{dl}(t-\Delta)\right|,-\sigma\ln\dfrac{2\sigma\beta}{b-a}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \sigma N\left(\ln\dfrac{2\sigma}{\alpha}-\ln(b-a)\right). \tag{10}$$

The decision device thus needs to compute the absolute distance between the uplink- and downlink pitch periods for all delays, Δ, of interest, saturate the absolute differences at $-\sigma\ln[2\sigma\beta / (b - a)]$, sum up the results and compare the sum with a threshold. The structure of the decision device is shown in Figure 2.
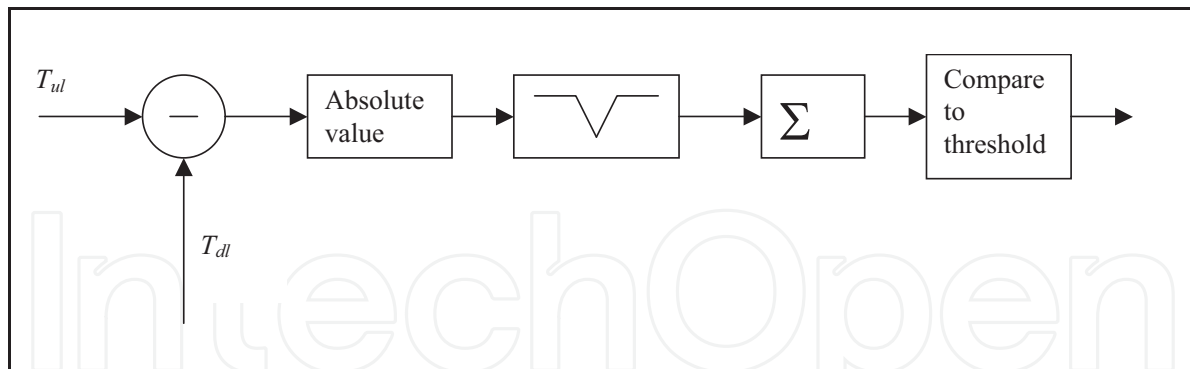
Fig. 2. Structure of the detector

4. Performance analysis

In this chapter we derive formulae for the probability of correct detection (we detect an echo when the echo is actually present) and the probability of false alarm (we detect an echo when there is none) of the detector. We start from reformulating the detector algorithm as

$$\frac{1}{N}\sum_{t=1}^{N} -\min\left(\left|w(t,\Delta)\right|,d\right) \overset{H_1}{\underset{H_0}{\gtrless}} -c, \tag{11}$$

where $c = \sigma\left[\ln(b-a) - \ln\frac{2\sigma}{\alpha}\right]$ and $d = -\sigma\ln\frac{2\sigma\beta}{b-a}$. The previous equation includes a nonlinearity that we denote as

$$y = h(w) = -\min\left(\left|w(t,\Delta)\right|,d\right). \tag{12}$$

$y=h(w)$ is a memoryless nonlinearity and hence the probability density function at the output of the nonlinearity is given by (Papoulis & Pillai 2002)

$$p_y(y) = \sum_{i=1}^{M} \left.\frac{p_w(w)}{\left|dy\middle/dw\right|}\right|_{w=w_i=h_i^{-1}(y)}, \tag{13}$$

where $p_w(w)$ is the probability density function of the input. M is the number of real roots of $y = h(w)$. That is, the inverse of $y = h(w)$ gives $w_1, w_2, \ldots, w_M$ for a single value of y. Note that in the problem at hand $M = 2$ and y is a piecewise linear function which has piecewise constant derivatives.

Let us first consider the case where the echo is present and, hence, the input probability density function is given by (3). We can see from (3) that the Laplace component is replaced by the uniform component in the probability density function in the points where

$$|w| = -\sigma\ln\frac{2\sigma\beta}{b-a} = d. \tag{14}$$

In addition we know from (12) that the output of the nonlinearity is saturated precisely at d. The probability density function of the output is therefore

$$p_y(y|H_1) = \frac{\alpha}{\sigma}\exp\left(-\frac{y}{\sigma}\right)[u(0)-u(d)] + \alpha\beta\frac{b-a-2d}{b-a}\delta(y-d), \tag{15}$$

where $u(\cdot)$ denotes the unit step function and $\delta(\cdot)$ denotes the Dirac delta function. The mathematical expectation of the output signal from the nonlinearity, $y|H_1$, is

$$E(y|H_1) = \alpha\left[\sigma - (\sigma+d)\exp\left(-\frac{d}{\sigma}\right) + \beta d\frac{b-a-2d}{b-a}\right], \quad \text{the second moment equals}$$

$$E(y^2|H_1) = \alpha\left[2\sigma^2 - (d^2+2d\sigma+2\sigma^2)\exp\left(-\frac{d}{\sigma}\right) - \beta d^2\frac{b-a-2d}{b-a}\right], \quad \text{and consequently the}$$

variance is $\sigma_{H_1}^2 = E(y^2|H_1) - E^2(y|H_1)$.

In the case the signal consist of contributions originating from the mobile side only (no echo) the input probability density is given by (7) and using this in (13) results in

$$p_y(y|H_0) = \frac{2}{b-a}[u(0-u(d))] + \frac{b-a-2d}{b-a}\delta(y-d). \tag{16}$$

The mean of this probability density function is $E(y|H_0) = 1 - \frac{d^2}{b-a}$, the second moment is

given by $E(y^2|H_0) = \frac{2}{3}\frac{d^3}{b-a} + \frac{b-a-2d}{b-a}d^2$ and the variance equals
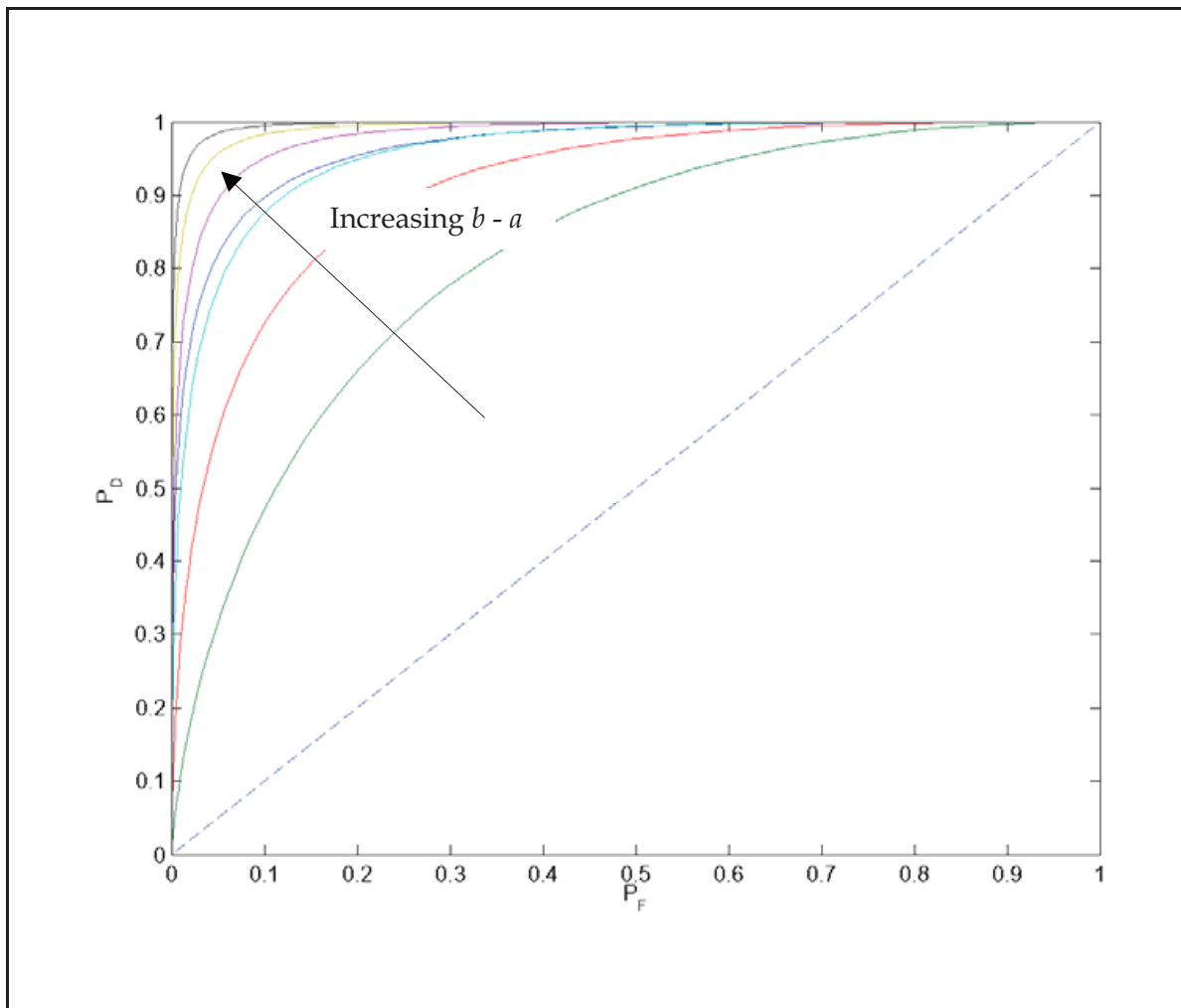
$\sigma_{H_0}^2 = E(y^2|H_0) - E^2(y|H_0)$.

Fig. 3. Receiver operating characteristics for varying *b - a*.

Note that our test variable in (11) is $\frac{1}{N}\sum_{t=1}^{N} -\min(|w(t-\Delta)|,d) = \frac{1}{N}\sum_{t=1}^{N} y_i$, which is a sample average of *N* i.i.d. random variables that appear at the output of the nonlinearity *y=h(w)*. According to the central limit theorem (Papoulis & Pillai 2002) the probability distribution function of such a sum approaches with increasing *N*, rapidly the Gaussian distribution with mean $E[y|H_i]$ and variance $\sigma_{H_i}^2 / N$ irrespective of the shape of the original distribution.

We can now evaluate the probability of correct detection (Van Trees 1971) as

$$P_D = \int_{-\infty}^{Th} p_y(y|H_1)dy = \frac{N}{\sqrt{2\pi}\sigma_{H_1}}\int_{-\infty}^{Th}\exp\left(-\frac{[y-E(y|H_1)]N}{2\sigma_{H_1}^2}\right)dy = \frac{1}{2}\left(\operatorname{erf}\left(\frac{[Th-E(y|H_1)]\sqrt{N}}{\sqrt{2}\sigma_{H_1}}\right)+1\right),$$

(17)

where *Th* is the threshold used in the test and erf denotes the error function $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt$. Correspondingly the probability of fault detection is

$$P_F = \int_{-\infty}^{Th} p_y(y|H_0) dy = \frac{N}{\sqrt{2\pi}\sigma_{H_0}} \int_{-\infty}^{Th} \exp\left(-\frac{[y - E(y|H_0)]N}{2\sigma_{H_0}^2}\right) dy = \frac{1}{2}\left(\text{erf}\left(\frac{[Th - E(y|H_0)]N}{\sqrt{2}\sigma_{H_0}}\right) + 1\right).$$

(18)

The Receiver Operating Characteristic (ROC), which is the probability of correct detection P_D as a function of probability of fault alarm P_F, is plotted in Fig. 3. The parameters used to compute ROCs are $N = 30$, $\sigma = 2$ and $\beta = 0.1$. The distance between the endpoints of the uniform density, b – a, varies from 4 to 12. One can see that the ROC curves approach the upper left corner of the figure with increasing b – a.

It is also important to note that the detector derived in this chapter has a robust behavior. The decision algorithm is piecewise linear in signal samples and in addition each entry of the incoming signal is saturated at d, meaning that no single noise entry no matter how big it is can influence the decision more than by d. Hence, the detector constitutes a robust test in the terminology used in (Huber 2004).

5. Practical considerations

The detector, as given by equation (10) is not very convenient for implementation, as it needs computation of a sum over all subframes with each new incoming subframe. To give formula (10) a more convenient, recursive, form we define a set of distance metrics D, one for each echo delay Δ of interest

$$D(t, \Delta) = tc - \sum_{i=1}^{t} \min(|T_{ul}(i) - T_{dl}(i - \Delta)|, d).$$

(19)

The distance metrics are functions of time t or more precisely the subframe number. Computation of the distance metric can now easily be reformulated as a running sum i.e. at any time t we compute the following distance metric for each of the delays of interest and compare it with zero

$$D(t, \Delta) = D(t - 1, \Delta) + c - \min(|T_{ul}(t) - T_{dl}(t - \Delta)|, d) \underset{H_0}{\overset{H_1}{\gtrless}} 0.$$

(20)

Note that a large distance metric means that there is a similarity between the uplink and downlink signals and the other way around, a small distance metric indicates that no similarity has been found. Also note, that one can easily introduce a forgetting factor into the recursive detector structure in order to gradually forget old data as it is customary in e.g. adaptive algorithms (Haykin 2002). We are, however, not going to do this here. The echo is detected if any of the distance metrics exceeds a certain threshold level, which is zero in this case. The echo delay corresponds to Δ with largest associated distance metric, $D(t, \Delta)$.

There are several practicalities that need to be added to the basic detector structure derived in Section 3:

1. Speech signals are non-stationary and there is no point in running the detector if the downlink speech is missing or has too low power to generate any echo. As a practical limit, the distance metric is updated only if the down-link signal power is above –30 dBm0.
2. By a similar reason there is a threshold on the down-link pitch gain. The threshold is set to 10000.
3. The detection is only performed on "good" uplink frames i.e. SID frames and corrupted frames are excluded.
4. It has been found in practice that $c = 7$ and $d = 9$ is a reasonable choice.
5. To allow fast detection of a spurious echo burst, the distance metrics are saturated at –200 i.e. we always have $D(t,\Delta) \geq -200$.

Additionally one can notice that the most common error in pitch estimation occurs at double of the actual pitch period. This can be exploited to enhance the detector. In the particular implementation this has been taken into account by adding a parallel channel to the detector where the downlink pitch period is compared to half of the uplink pitch period

$$D(t,\Delta) = D(t-1,\Delta) + c_1 - \min\left(\left|\frac{T_{ul}(t)}{2} - T_{dl}(t-\Delta)\right|, d_1\right) \mathop{\gtrless}_{H_0}^{H_1} 0, \qquad (21)$$

where the constants c_1 and d_1 are selected to be smaller than the corresponding constants c and d in (20) to give a lower weight to the error channel as compared to the main channel. Only one of the updates given by (20) and (21) is used each time t. The selected update is the one that results in a larger increase of the distance metric.

6. Simulation results

Our simulation study is carried out with the aim of investigating how well the derived detector works with speech signals. We first investigate if the distribution adopted in this work can be justified. This is followed by some experiments clarifying detection performance of the proposed algorithm using recordings made in an actual mobile network and finally we investigate the resistance of the detector to disturbances.

6.1 Distribution of pitch estimation errors
In this section we investigate the distribution function of the pitch estimation errors. The main question to answer is if the distribution function adopted in Section 3 is in accordance with what can be observed in the simulations.
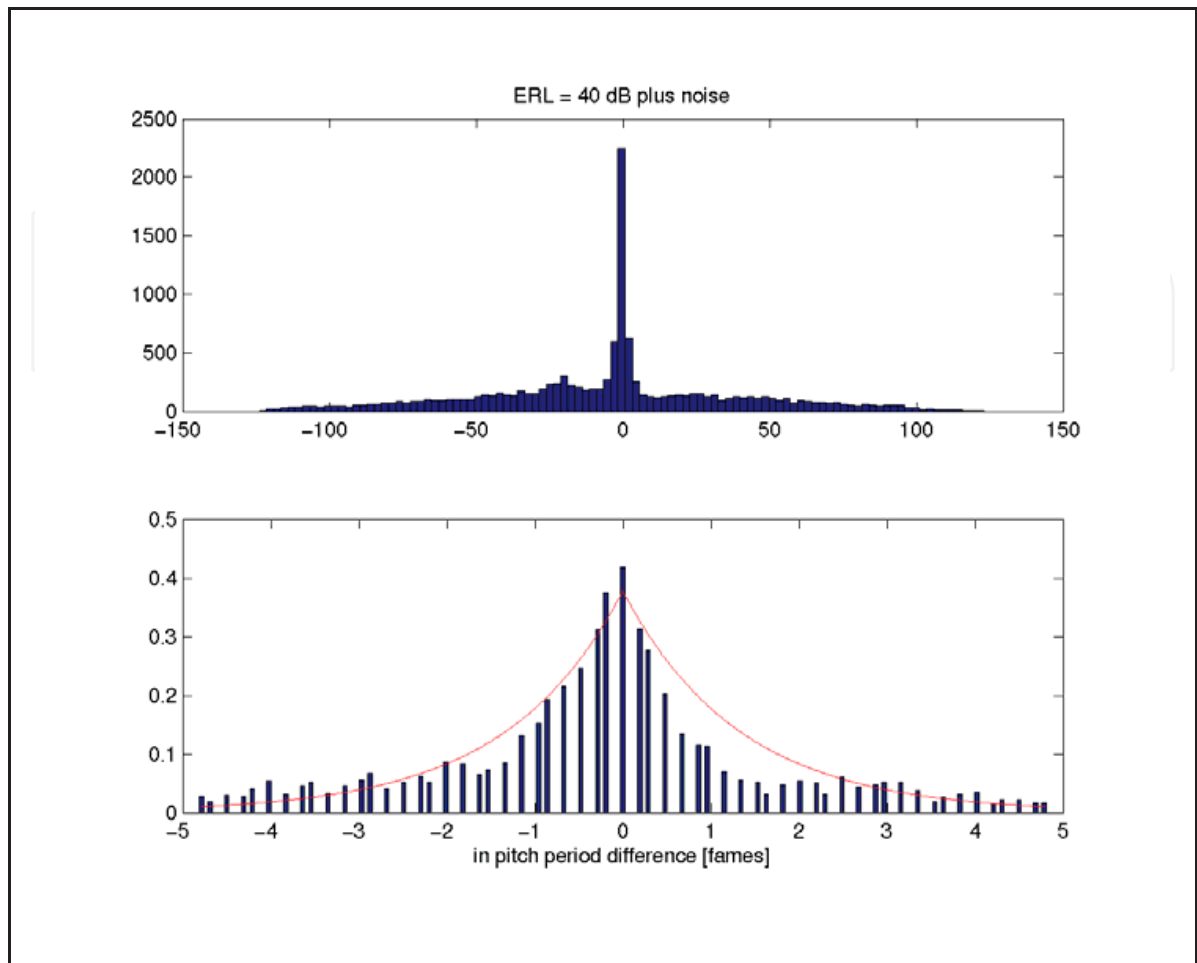
Fig.4. Histogram of pitch estimation errors. Echo path: single reflection and IRS filter, ERL = -40dB. Near end noise at -60 dBm0

To answer this question a two minute long speech file that includes both male and female voices at various levels was first coded with the AMR12.2 kbit/s mode codec and then decoded. Then a simple echo path model consisting either of a single reflection or the IRS filter (ITU-T G.191) was applied to the signal and the signal was coded again. Echo return loss was varied between 30 and 40 dB. The estimated pitch was registered from both codecs and compared. The pitch estimates were used only if the downlink power was above –40 dBm0 for the particular frame. A typical example is shown in Figure 4. The upper plot shows the histogram of pitch estimation errors. A narrow peak can be observed around zero and the histogram has long tails ranging from –125 to 125 (which are the limiting values for differences between two pitch periods). The lower plot shows the Laplace probability density function fitted to the middle part of the histograms. One can see that there is a reasonable fit.

6.2 Detection performance
Recordings made with various mobile phones were used to examine the detection performance. All the distance metrics in (20) were initialized to -50 and the echo was declared to be present if at least one of the distance metrics became larger than zero. Validity

of the detection was verified by listening to the recorded file and comparing the listening and detection results. The two were found to be in a good agreement with each other.
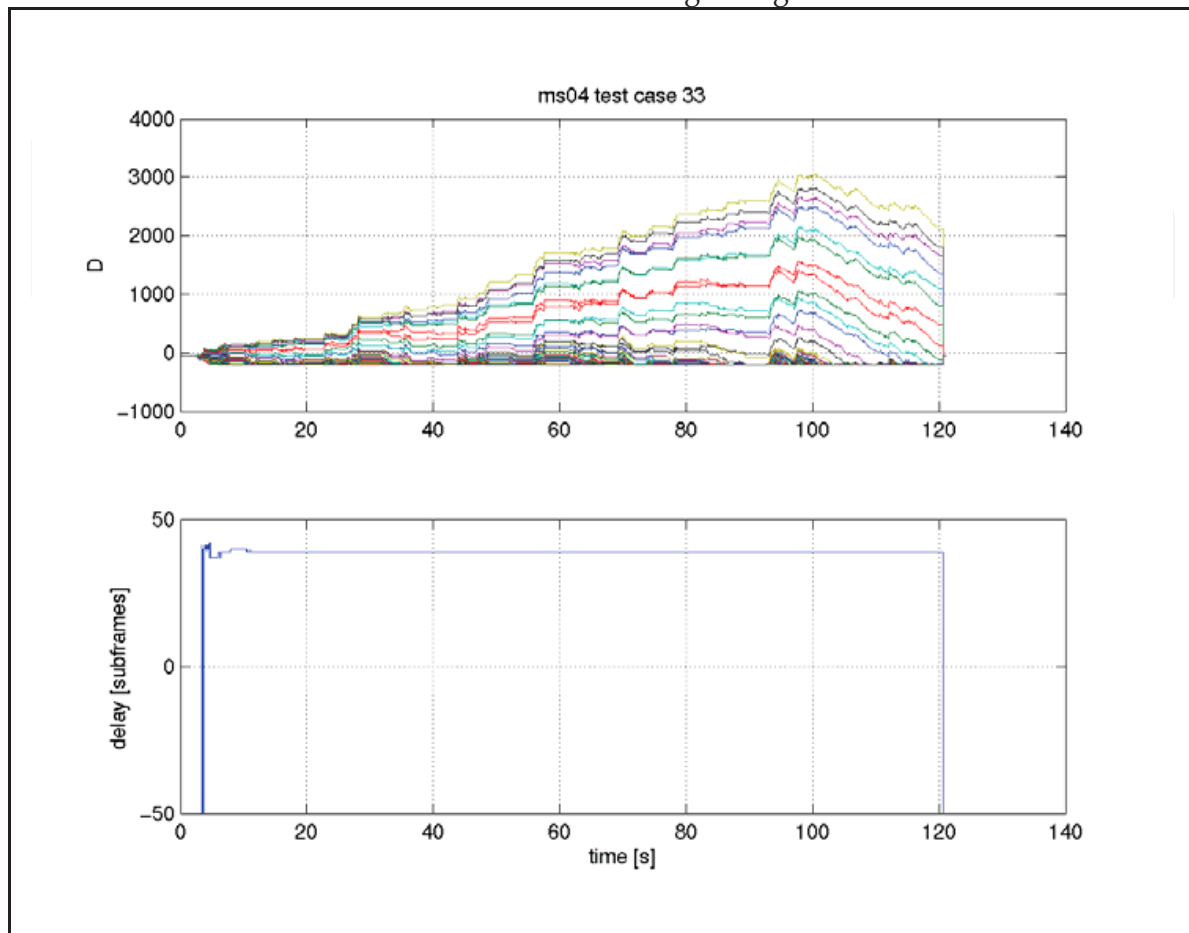

Fig. 5. Distance metrics (upper plot) and estimated delay (lower plot)

The delay was estimated as the one corresponding to the largest distance metric. As the experiments were done with signals recorded in real mobile systems, the author lacks knowledge of the true echo delays in the test cases. However, the estimates were proven in practice to provide good enough estimates for usage in a practical echo removal device.
Let us finally note that the resolution of the delay estimate is 5 ms due to the 5 ms subframe structure of the AMR speech codec.
A typical case with a mobile that produces echo is shown Figure 5. One can see that in this example echo is detected and the delay estimate stabilizes after a couple of seconds to 165 ms, which is a reasonable echo delay for a GSM system.

6.3 Resistance to disturbances
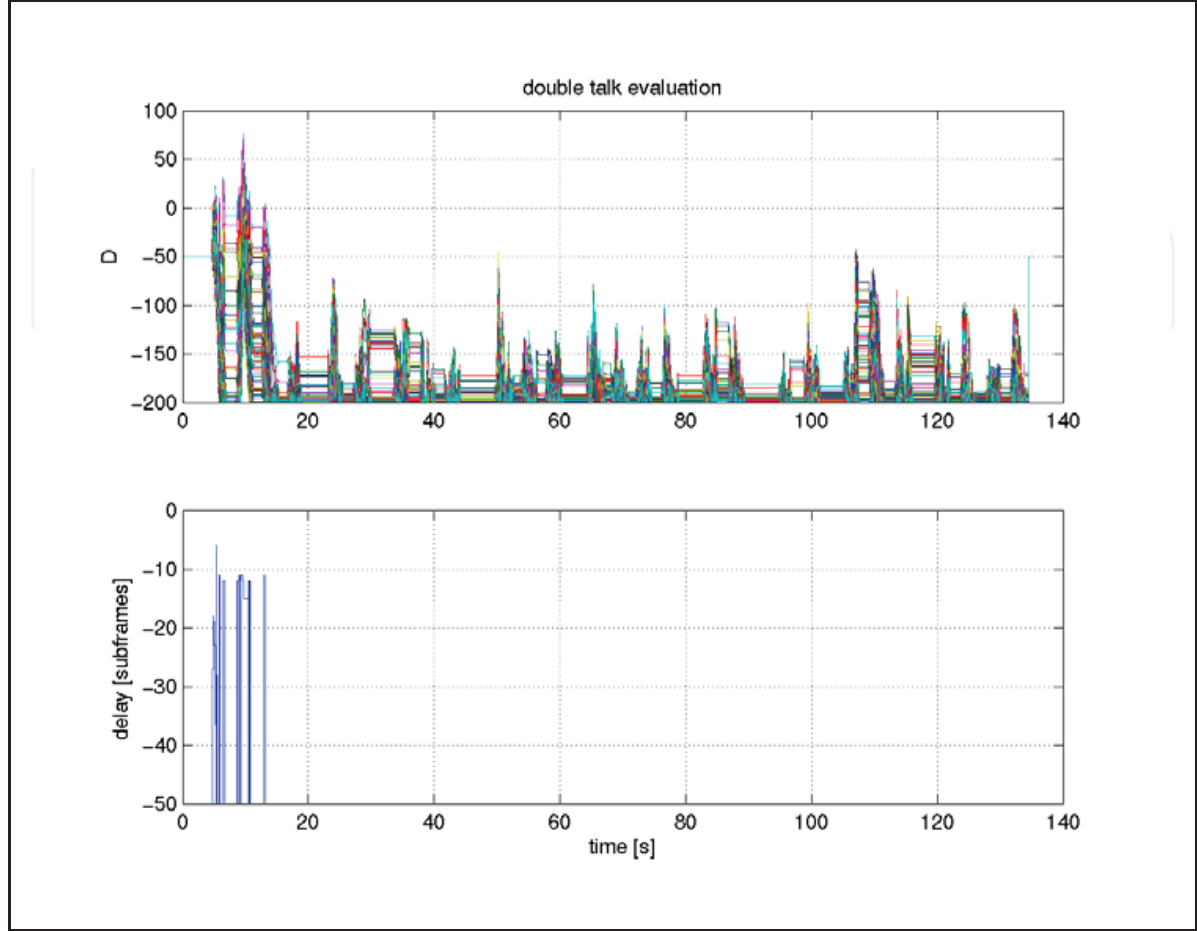
Fig. 6. Evaluation of the detector in severe double talk conditions, mobile side male, network side female.

Finally we check that the detector is not unnecessarily sensitive to disturbances. In echo context the most common disturbance is so called double talk i.e. the situation when both parties of the telephone conversation are talking at the same time. In this situation speech from the near end side forms a strong disturbance to any algorithm that needs to cope with echoes. The proposed detector algorithm was verified in a large number of simulations involving speech signals from both sides of the connection and its double talk performance was found to be good. This is expected as robustness against disturbances is built in the algorithm in form of limited distance measure update in equation (20).

As another and perhaps somewhat spectacular demonstration of double talk performance, we used two speech files with male and female voices speaking exactly the same sentences the same time. The result is shown in Figure 6 with female voice from network side and male voice from mobile side. In this case there are some fault echo detections initially, partly caused by initialization of the distance metrics to –50. Duration of the fault detection is, however, limited to the first two sentences (14 seconds) of the double talk. There was no echo detected in the opposite scenario i.e. male voice from the mobile side and female voice from the network side. Taking into account that it is very unlikely that in an actual

telephone call both sides would talk the same sentences simultaneously we conclude that the detector is reasonably resistant to double talk.

7. Conclusion

This chapter deals with the problem of processing AMR coded speech signals without decoding them first. Importance of such algorithms arises from the fact that not all calls need enhancement and even if they need the quality loss from decoding and coding speech again may be higher than the potential improvement due to speech enhancement. Processing the signals directly in the coded domain avoids this problem.

The chapter proposes a detector that can be used to detect presence of echo generated by mobile phones and estimate its delay. The detector uses saturated absolute distance between uplink and downlink pitch periods as a similarity measure and is hence a robust algorithm. Performance of the detector was analysed and the equations for detection and error probabilities were derived. Finally a good performance of the detector with real life signals was demonstrated in our simulation study.

8. References

3GPP TS 26.090 V6.0.0 (2004-12) 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Mandatory Speech Codec speech processing functions; *Adaptive Multi-Rate (AMR) speech codec; Transcoding functions (Release 6).*

Haykin, S. (2002) *Adaptive Filter Theory, fourth edition,* Prentice Hall

Huber, P. (2004) *Robust Statistics*, Wiley & Sons

ITU-T Recommendation G.160 (2008) *Voice Enhancement devices*

ITU-T Recommendation G.191, (2005) *Software Tools for Speech and Audio Coding Standardization*

Kay, S. (1998) *Fundamentals of Statistical Signal Processing, Volume II, Detection Theory*, Prentice Hall

Papoulis, A. and Pillai, S. U. (2002) *Probability, Random Variables and Stochastic Processes, Fourth edition,* McGraw Hill

Perry, A. (2007) *Fundamentals of Voice-Quality Engineering in Wireless Networks*, Cambridge University Press

Signal Processing, Volume 86, Issue 6, June 2006, Special issue on Applied Speech and Audio Processing.

Sondhi, M and Berkley, D. (1980) Silencing Echoes in Telephone Network, *Proc. IEEE,* Vol. 68, No. 8, Aug. 1980 pp. 948-963.

Van Trees, H. L. (1971) *Detection, Estimation, and Modulation Theory*, Wiley & Sons
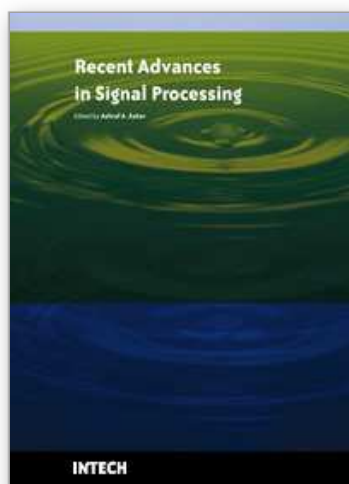
Recent Advances in Signal Processing

Edited by Ashraf A Zaher

ISBN 978-953-307-002-5

Hard cover, 544 pages

Publisher InTech

Published online 01, November, 2009

Published in print edition November, 2009

The signal processing task is a very critical issue in the majority of new technological inventions and challenges in a variety of applications in both science and engineering fields. Classical signal processing techniques have largely worked with mathematical models that are linear, local, stationary, and Gaussian. They have always favored closed-form tractability over real-world accuracy. These constraints were imposed by the lack of powerful computing tools. During the last few decades, signal processing theories, developments, and applications have matured rapidly and now include tools from many areas of mathematics, computer science, physics, and engineering. This book is targeted primarily toward both students and researchers who want to be exposed to a wide variety of signal processing techniques and algorithms. It includes 27 chapters that can be categorized into five different areas depending on the application at hand. These five categories are ordered to address image processing, speech processing, communication systems, time-series analysis, and educational packages respectively. The book has the advantage of providing a collection of applications that are completely independent and self-contained; thus, the interested reader can choose any chapter and skip to another without losing continuity.

INTECH

open science | open minds