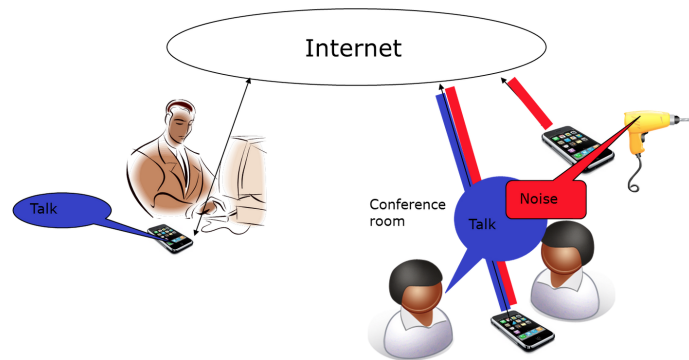KUNGLIGA TEKNISKA HÖGSKOLAN

Project Report

Noise and echo cancellation in a teleconference

Authors:
Animesh DAS
Jonas SEDIN
Mohammad ABDULLA
Thomas GAUDY
Xavier BUSH

Advisor:
Per ZETTERBERG

Spring 2015

Contents

Chapter 1

Background

1.1 Introduction of noisy environments

It is a fact that the scenarios with phone calls involved are increasing every day. This situation implies an increase of the probability of being in a noisy scenario, specially in big cities. As a result of the discomfort that the users suffer in these noisy environments, engineering and science have worked with different approaches to solve this problem.

The diversity of noise nature and its sources lead the engineering to a big challenge: develop high performance soltions in these diverse environments. When facing noise cancellation is very important to take into account the variability that the noise may experience, as previously said. Duration of the noise sequences (from ms to long sequences), color of the noise and stationarity are possible classifications of the noise and each classification implies different ways of treating it. Therefore, a lot of systems are using combined techniques to reach the best possible performance, which has been naturally the case of this project.

1.2 Historical Overview

Before presenting the proposed solutions and approaches of the project, it is needed a historical overview to understand how have the group been influenced and which have been the patterns of research.

1.3 Description of the project

The problem proposed by the course $EQ2440$ has been a "Noise and echo cancellation of a teleconference". The general scenario is that the first of the two speakers of the teleconference is in a noisy environment and the clear goal is to cancel as much noise as possible in order that the second speaker could receive a cleaner speech and make the conversation more comfortable. As said in **??**, there are different approaches to solve this problem, where several of them require the availability of pure noise recordings, in our case recorded with a third phone placed close to the noise source. To have a clearer overview of the scenario the Figure 1.1 shows an approximate scheme easy to understand.

When talking about denoising a teleconference there are two factors to take into account, techniques to cancel the noise and the possibility of their implementation in a real time application. The real time application has been, as expected, a big challenge because

it implies good performance in terms of cancellation with the minimum reachable delay to conserve the naturalness of the conversation.
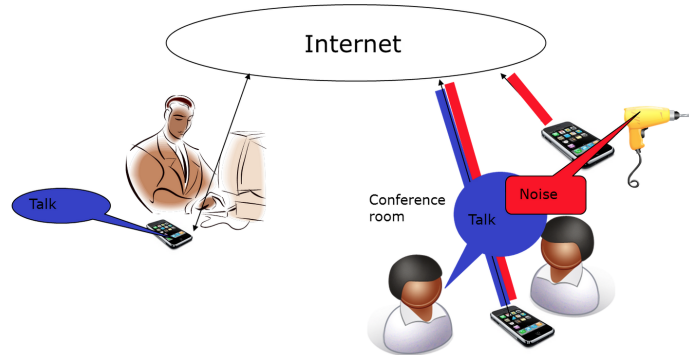

Figure 1.1: Scenario to solve

1.4   Goal

As commented in 1.3, the goal is to cancel the noise contribution in the conversation between the two speakers of the teleconferences. With the purpose to simplify the scenario, it will be assumed that only one of the speakers is surrounded by noise and the main noise source is known as well.

As in every engineering project, the group had to find a compromise between performance in noise cancellation and viability of implementation in real life. As it will be explained in 2, the computational cost is a big constrain and the best performance of certain approaches (3) introduce too much delay because of this reason. As a consequence, not always the best solution will be possible to implement in the real time version of the project.

As a contrast, the personal goals of the project members are to learn form the teamwork environment, learn a research methodology, research criteria and certain skills of management that might be used in the performance of a Master Thesis (as an inmidate future) and in a research or business environment.

The new knowledge acquisition is obviously another personal goal of all the team members.

1.5   Organizationn and Human Resources

The organization of the project consists in electrical engineering students at different stages of the studies and within different specializations. In order to make the team as efficient as possible, the project has been divided in four different groups: *Theory Group*, *Android Group*, *Multimedia Group* and *Management Group*, all of them explained in detail in 2.

The distribution of the team members has been as follows.

- Animesh Das

    - Role: Management Group
    - e-mail:animeshu1989@gmail.com (animeshd@kth.se)

- Telephone: +46 737155575

- Jonas Sedin

  - Role: Theory Group & Android Group
  - e-mail: sedinjo@gmail.com (jonassed@kth.se)
  - Telephone: +46 704252951

- Mohammad Abdulla

  - Role: Android Group & Multimedia Group
  - e-mail: hamodiilatch@gmail.com (mabdulla@kth.se)
  - Telephone: +46 737393276

- Thomas Gaudy

  - Role: Android Group
  - e-mail: gaudy.thomas@gmail.com (gaudy@kth.se)
  - Telephone: +46 760936034

- Xavier Bush

  - Role: Theory Group & Management Group (Project Leader)
  - e-mail: xavier.bush@gmail.com (xbush@kth.se)
  - Telephone: +46 764141834

The sponsor members as Project Examiner/Supervisor and Project Support are:

- Per Zetterberg

  - Role: Project Examiner
  - e-mail: perz@ee.kth.se
  - Telephone: +46 8 790 77 85

- Hadi Ghauch

  - Role: Group Assistant
  - e-mail: ghauch@kth.se

- Martin Ohlsson

  - Role: Android Guru
  - e-mail: martinoh@kth.se
  - Telephone: +46 87907818

Chapter 2

Methodology

This chapter shows the methodology that the group has followed since the projecte started. On the first hand, it goes without saying that the project group has followed the *Scientific Method* in the implementation of the project. On the second hand, as commented in 1.5, the group has been divided in three groups explained in the following subsections.

2.1 Theory Group

The *Theory Group* had as its main goal finding solutions to cancel the present noise in the teleconference. Nevertheless, a constraint of the group has been the computational cost that the implementation have, where all the details may be found in 3.

The fact that three members of the project had recent and good background in Adaptive Signal Processing, which has been one of the chosen approaches to face the noise cancellation, made easier the making of the groups. Moreover, the *Theory Group* avoided the first stage in theory research, which is the most difficult part when starting a new project.

In terms of methodology, the *Theory Group* has followed next steps:

- Make research in suitable algorithms.

- Record with the given mobile phones both signals: 'voice+noise' and 'noise'.

- Test the performance in MATLAB.

- Check the computational cost in MATLAB.

- Check the possibility to transfer the solutions from MATLAB to Android.

Because of the presence of Jonas Sedin in the *Theory Group* and the *Android Group*, it has been possible to design theory solutions think in the availability to transfer them to Android coding.

2.2 Android Group

This group has had two different duties:

- Android tasks: these tasks aimed to be an introdution to the Android programming.

9

- Noise cancellation in Android: has been the transfer from the proposed solution in the *Theory Group* to Android.

- Real time application: optimize the code to decrease the computational cost and make possible a real time application.

In this case, even if Thomas Gaudy has not been a part of the *Theory Group*, his background in Adaptive Signal Processing has helped in the implementation of the proposed solutions.

2.3   Multimedia Group

This group is only formed by Mohammad Abdulla who has taken care of all the technical preparation of the presentations of the project:

- Video of the project: explanation of the project with real examples.

- Power Point Presentation: power point to use in the Grand Final Presentation.

- Bloopers Video: video containing bloopers and funny moments during the performance of the project.

2.4   Management Group

The *Management Group* has taken car of the drawing up of those tasks that were oriented to plan, follow up and present the work of the project.

The main documents are listed below:

- Project Plan: document that cointained a brief description of the project, time resources and planning and tasks planning.

- Progress Report: document containing the follow up of the projects in terms of achievements and resources. It is a updated version of the Project Plan. The last Progress Report can be found in 6.

- Project Report: this precise document. It contains the description of all the project in all the areas.

Besides the main documents, the *Management Group* has taken care of next tasks:

- Meetings: there has been no need to schedule meetings because of the small size of the group and the cross-group duties.

- Scheduling: the planning of all the tasks was designed by this group and it has been properly followed.

2.5   Cross-Groups Duties

In terms of making easier the transfer of information between groups, the makeup of the groups has followed a cross-specialization criteria. Therefore, there are team members in more than one group, and it has been possible due to the wide scope of skills that the team members have (1.5).

With this makeup, there have

Chapter 3

Theory

The *Theory Group* has been working within different scenarios but mainly has followed what is called *Adaptive Signal Processing*, as it has been explained in **??**. Some of them have ended in good performance algorithms where only two of them have been transferred to Android and some other have not reached the wanted outcome.

In terms of goals, it has been very important for the group to reach a good noise cancellation and preserving as much as possible the voice quality.

3.1 Successful Approaches

3.1.1 LMS Algorithm

Introduction

The LMS (Least Mean Square) Algorithm is an adaptive filtering technique used to estimate the unwanted component in the signal that wants to be filtered and, a posteriori, subtract it to "clean" the wanted signal.

To use this technique it is a needed to have both signals, the one that contains the useful informationn($y(n)$) and the noisy one ($x(n)$).

The process used in this technique is a recursive calculation samples by sample were N previous taps of the wanted signal are used to calculate the current filter sample, used later on for the subtraction.

The original problem to solve in LMS is shown in the Equation 3.1, where is clearly visible that the LMS minimizes the MSE using the instantaneous error.

$$\hat{\theta}(n) = \hat{\theta}(n-1) - \frac{\mu}{2}) \frac{\partial}{\partial \theta} MSE(n, \theta)|_{\theta = \theta(\hat{n-1})} \tag{3.1}$$

Where after the appropriate calculations, the final version ready to implement in MATLAB is next:

$$\hat{\theta}(n) = \hat{\theta}(n-1) + \mu Y(n)(x(n) - Y^T(n)\hat{\theta}(n-1)) \tag{3.2}$$

In the Equation 3.2:

- $\hat{\theta(n)}$: estimation of the current sample of the filter

- $\hat{\theta(n)}$: estimation of the previous sample of the filter

- $Y(n)$: vector that contains the N taps used in the implementation

- $x(n)$: sample of the unwanted signal

In terms of cost operation, LMS is very efficient thus only needs O(N) operations in each iteration. A priori, this is feasible for a real time implementation and this one of the reason that made us start with this approach.

Implementation & Results

The implementation used in MATLAB is based on the one that Jonas, Thomas and Xavier have done during the course of *Adaptive Signal Processing (EQ2400)*.

Once all the recordings where done, the first tests of the LMS gave different performances depending on the number of previous taps used to calculate the filter.

- $N = 1000$: very good performance with too much delay

- $N = 100$: good performance

Another very important parameter to consider is the *Step Size* μ. The value of μ has a direct impact on the stability and convergence of the algorithm. The general rule of choosing the right *step size* is shown in **??**

$$0 < \mu < \frac{2}{\lambda_1} \tag{3.3}$$

where λ_1 is the maximu eigen value of $I - \mu\Sigma_{YY}$.

Even though there is an upper limit for the value of μ, there is not a lower limit (besides 0). Therefore, the group followed a test-and-set empirical criteria where different μ values where tested and the value who gave the best auditive performance was the chosen. In the last case, the best performance was given by $\mu = 5 \cdot 10^{-5}$.

These results are reasonable: it exists a linear dependency of the performance with the number of previous samples used to calculate the current one.

To show some graphic results of the performance of this algorithm, on the one hand Figure 3.2 and **??** show the time and frequency response of the filtered sequence with $N = 100$ taps in comparison with the original recording. There is some audible noise reduction that in the figure can be seen in the noisy frames where there is more presence of noise in the original plot than in the filtered one. When it comes to the frequency domain, we can appreciate that there are several small peaks at low frequencies and the contribution of the peak at $f = 0.6$ (where $fs = 44.1$ KHz) has been slightly decreased.

On the other hand, the Figures 3.4 and **??** show the same comparison as before but with a filter of $N = 1000$ taps. In this case the Figure 3.4 shows that the noise has less presence. The drawback comes with a certain distortion in the audio file that can be seen in the Figure **??** as a modification in the frequency domain with new peaks in a coarser spectral response at lower frequencies.

In terms of computational time, in MATLAB the LMS with $N = 100$ taps takes 2.22 seconds, while the LMS with $N = 1000$ takes 3.78 seconds (increasing of a 70%). This does not make a huge difference in MATLAB but in a real time application it might, therefore it will be treated by the *Android Group*.
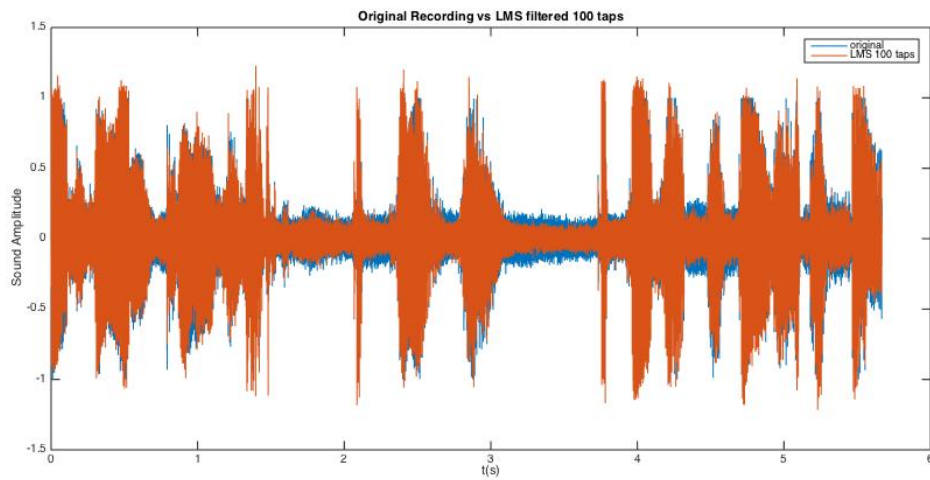
Figure 3.1: Time Domain - LMS result with N=100

3.1.2 Wiener Non-Causal Filtering

3.1.3 Speech Enhancement Using a-Minimum Mean- Square Error Short-Time Spectral Amplitude Estimator

3.2 Unsuccessful Approaches

3.2.1 Kalman Filtering

3.2.2 RLS
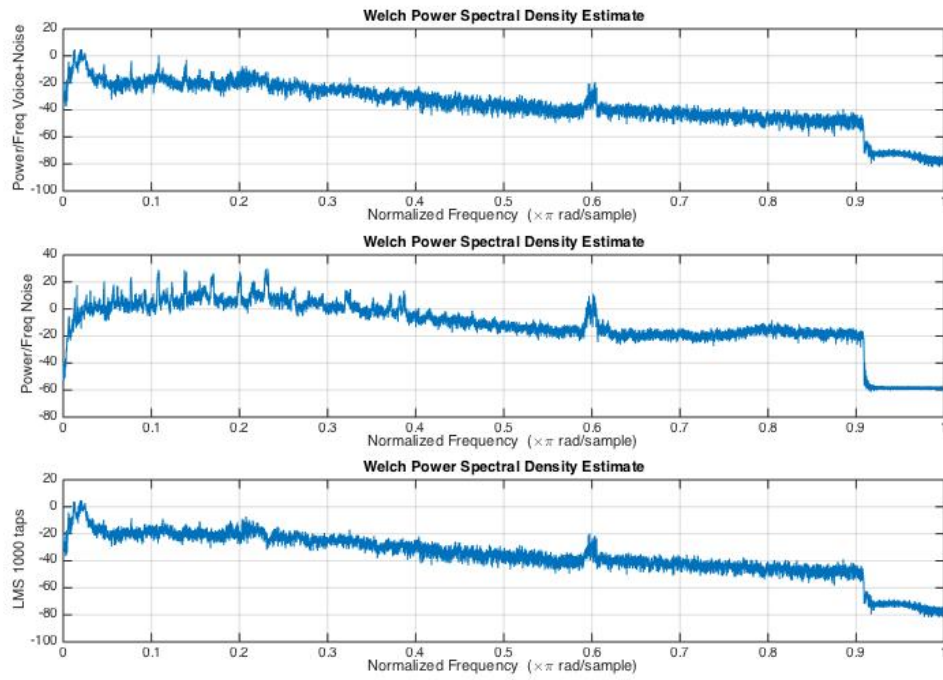
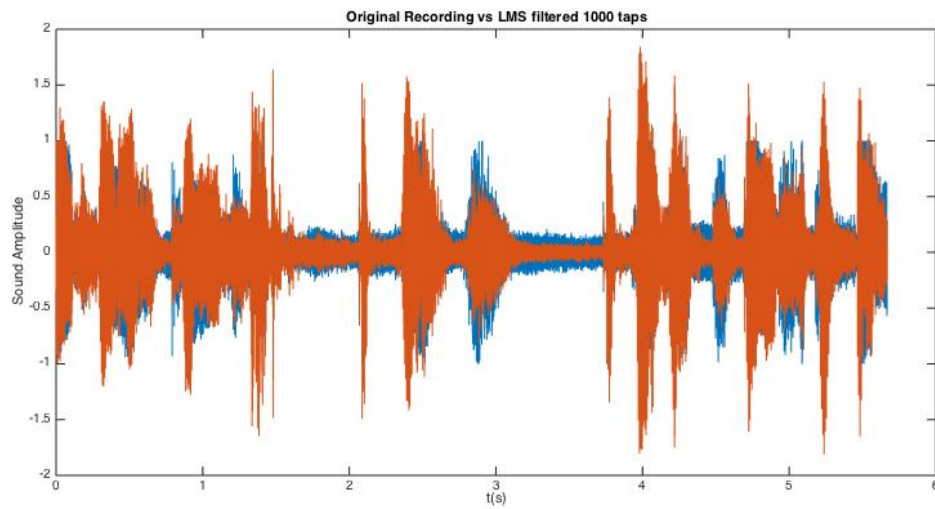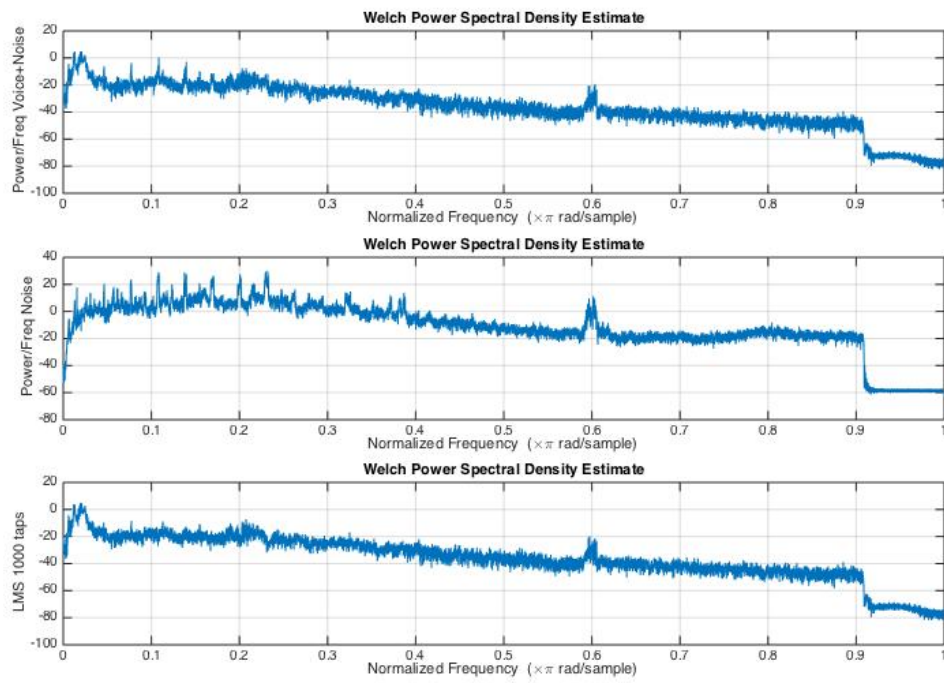Figure 3.2: Frequency Domain - LMS result with N=100

Figure 3.3: Time Domain - LMS result with N=1000

Figure 3.4: Frequency Domain - LMS result with N=1000

Chapter 4

Android

Chapter 5

Conclusions

Chapter 6

Appendices