**République Algérienne Démocratique et Populaire**
**Ministère de l'Enseignement Supérieur et de la Recherche Scientifique**

**Ecole Nationale Polytechnique**

**Département : ELECTRONIQUE**

**Laboratoire Signal &Communications**

## Thèse de Doctorat

**Présenté Par :**
**AHFIR Maâmar**

# DEREVERBERATION SYSTEMS DESIGN FOR ROOM ACOUSTICS

**Devant le jury**

| | |
|---|---|
| **Président** | Prof M.S.Boucherit - ENP Alger |
| **Examinateur** | Prof A.Guessoum - Université de Blida |
| **Examinateur** | Prof M.Bensebti - Université de Blida |
| **Examinateur** | Prof I.Kale -University of Westminster (London-UK) |
| **Examinateur** | Dr H. Bousbia - ENP Alger |
| **Rapporteur** | Prof D.Berkani - ENP Alger |

## Mars 2008

**E.N.P. 10, Avenue Hassen Badi –El-Harrach– ALGER**

**ملخص**

إن تقنيات التعديل الجزئي جد مهمة لأنها تسمح لنا في نفس الوقت بتقليص تأثير الرفربريشن و كلفة تطبيق المرشحات العكسية من نوع **FIR** - الاستجابة المنتهية للومضة -.
في هذه الأطروحة نقترح تقنيتين اثنتين لتصميم أنظمة درفربريشن على أساس التعديل الجزئي من اجل معالجة الصوتيات داخل القاعات.
طريقة هومورفيك استعملت من اجل تصميم المرشحات العكسية ذات الطور الأدنى و مستقرة التعديل لاستجابات الومضات الصوتية ذات الطور لا أدنى لقاعة صغيرة ثم لأخرى كبيرة.
في التقنية الأولى نقوم بتغيير بعض الأقطاب المهيمنة في دالة التحويل للمرشح العكسي بأقطاب جديدة قبل القيام بعملية **DFT**العكسية – تحويلية فوريه المتقطعة- .
النتائج المحصل عليها بالنسبة لاستجابة الومضة الصوتية المقاسة داخل سيارة صغيرة تبين انه باستعمال التقنية المقترحة نستطيع مراقبة جودة الصوت بدقة, أحسن من استعمال التقنية التقليدية.
في التقنية الثانية نقوم بتطبيق عملية صقل تكرارية لاستجابة الومضة الأصلية الممثلة لقاعة كبيرة قبل إجراء العملية العكسية بعد هذا نحصل على استجابات الومضات مقلصة في الزمن لكن مطابقة لمبادئ الإدراك السمعي لدى الإنسان. هذه الاستجابات المقلصة في الزمن تستعمل من اجل تصميم مرشحات التعديل.
نتائج المحاكاة المحصل عليها بالنسبة لقاعة محاضرات تسودها رفربريشن تم تأكيدها بإجراء اختبارات السمع على جهاز **DSP** لتكساس انسترمنت.



**ABSTRACT**

Partial equalization techniques are very interesting because they simultaneously allow reduction of **reverberation** effect and implementation complexity of FIR inverse filters. In this work, two different techniques for **dereverberation** systems design based on partial equalization for room acoustics have been proposed. The Homomorphic method is used to design stable minimum-phase inverse filters (equalizers) for non-minimum-phase acoustic impulse responses corresponding to small and large rooms. In the first proposed approach, some of the dominant poles of the inverse filter transfer function are replaced by new ones before carrying out the inverse DFT. Results for an impulse response measured in the car interior show that by using the modified version we can control the sound quality more precisely than when using the standard method. In the second approach, an iterative simple smoothing is applied to the original impulse response of a large room before its inversion. Corresponding time reduced impulse responses are derived which conform to perceptual principles. The smoothed impulse responses are then used to design equalization filters. Simulation results for a reverberant audio-conferencing room have been validated using listening tests on Texas Instruments DSP board.



**RESUMEE**

Les techniques d`égalisation partielle sont très intéressantes parce qu`elles permettent a la fois de réduire l`effet de la **réverbération** et la taille des filtres inverses RIF pour l`implémentation. Dans ce travail deux techniques différentes de conception des systèmes de **déréverbération** basées sur une égalisation partielle, pour l`acoustique des salles ont été proposées. La méthode homomorphique a été utilisée pour calculer des filtres inverses a phase minimale stables RIF (égaliseurs), pour l`égalisation des réponses impulsionnelles acoustiques a phase non minimales, dans les cas de petite et grande salle. Dans la première approche, quelques pôles dominant de la fonction de transfert du filtre inverse sont remplacées par des nouveaux pôles avant d`effectuer la TFD inverse. Les résultats obtenus dans le cas d`une réponse impulsionnelle mesurée dans un habitacle d`une voiture montrent bien qu`on peut mieux contrôler et avec précision la qualité du son en utilisant la version modifiée que d`utiliser la méthode standard. Dans la deuxième approche, un lissage itératif simple est appliqué a la réponse impulsionnelle originale d`une grande salle avant son inversion. Des réponses impulsionnelles de taille réduites sont générées mais qui sont conformes aux principes de la perception. Ces réponses sont donc inversées pour construire des filtres d`égalisation. Les résultats de simulations dans le contexte d`une salle d`audioconférence réverbérante ont été validés par les tests d`écoute sur un kit de DSP de Texas Instrument.

DEDICATION

To

The memory of my father, my grand father and grand mother.

To my mother,

my wife, my daughter Marwa, my sons Amine and Abderaouf,

all my family in particular Cherif and Abdelkader.

And also to

my close friends of Landon, Cherchell and Laghouat : M, F, H , R, N, K A, B and S,

who made all the possible, for their endless encouragement and patience.

# ACKNOWLEDGEMENTS

I would like to thank

Dr Benallal Ahmad who has proposed this project;

Prof Berkani Daoud for his supervision in Algeria and his encouragement to realize this project;

Prof Izzet Kale who has given me the opportunity to use the facilities of the applied DSP and VLSI research group laboratory at Department of Electronic Systems of the University of Westminster in London (UK) and for his supervision during the realization of this project;

Dr Artur Krukowski for his contribution to this project;

Dr Ediz Cetin for his helps and advises in writing the thesis and his graduation students for their participation to listening tests evaluation;

The Chairman of Examiners Prof M.S.Boucherit, Prof A.Guessoum, Prof M.Benssebti and Dr H. Bousbia to have accepted my thesis examination.

I would also like to thank

The Department of Electronic Systems of the University of Westminster to have accepted me as a visiting researcher since March 2004;

And the CNET–Lanion (France) for the entire recorded wave files in anechoic chamber and measured acoustic impulse responses.

# LIST OF FIGURES

# LIST OF TABLES

## CONTENTS



### CHAPTER 1
### ACOUSTIC IMPULSE RESPONSES MEASUREMENTS



### CHAPTER 2
### PARTIAL EQUALIZATION FOR SMALL ROOMS



### CHAPTER 3
### PARTIAL DEREVERBERATION FOR LARGE ROOMS



### CHAPTER 4
### RESULTS VALIDATION

**M. Ahfir**, I. Kale, A. Krukowski and D. Berkani, "Partial Equalization of Non Minimum-Phase Impulse Responses", EURASIP Journal on Applied Signal Processing, Volume 2006, n°13, ISSN 1110-8657.

**M. Ahfir**, I. Kale, A. Krukowski and D. Berkani, "Room Equalization Based on Iterative Simple Complex Smoothing of Acoustic Impulse Responses", 12° IEEE International Conference ICECS'05, TUNISIA, Vol 1, pp 256-259, Dec 2005.

**M. Ahfir,** I. Kale, D. Berkani and E. Cetin, "Implementation of a Reduced Complexity FIR inverse filter on DSP Board for Room Equalization", 2008 IEEE International Symposium on Industrial Electronics ISIE 08, Accepted paper, Cambridge, UK, July 2008. .

# INTRODUCTION

Reverberation phenomenon is caused by the effect of the room where the useful sound source is placed. This is because of acoustic wave reflections. Reverberation is very sensitive in communications with Hands-free systems and in general, when a sound recording is carried out in an enclosure without a particular acoustical treatment. When reverberation is relatively significant, the source seems to be subjectively far away to the listener and then the speech intelligibility is degraded (barrel-like quality).

Dereverberation is to make the desired signal completely or partially clean of the effect of the room. A perfect dereverberation would leave only the wave coming directly from the source and would produce the same effect than a sound recording taken in an anechoic chamber. This is not required in room acoustics because a moderate room effect can make the speech to sound natural.

A first approach of dereverberation is an inversion of the acoustic channel represented by its measured impulse response, which provides the room effect between the source and the microphone installed in the room. This approach is not realistic because the frequency response of the room has many zeros which rapidly vary with the microphone position in the room. Therefore, it is required to use methods taking into account the temporal and spatial aspects of the reverberation phenomenon [1].

Some methods based on direct inversion of the acoustic channel have been proposed by some authors [2, 3, 4], but that still need modifications to be adapted to corresponding contexts, to better respond to the desired speech quality. In this thesis, we propose two equalization techniques (inversion methods) of the acoustic channel. The first technique is recommended in the context of small room (car, telephone-booth, etc.), and the second is recommended in the context of large room, such as audio-conferencing and video-conferencing rooms.

The general idea is to realize stable equalizer Finite Impulse Response (FIR) digital filters (FIR inverse filters) taking into account the temporal and spatial aspects of the reverberation phenomenon in both cases, small and large rooms.

In the first case a partial equalization of the impulse response representing this context has been realized. This technique based on homomorphic method, allows controlling the speech quality with more precision [5]. In the second case where reverberation is more significant, an iterative smoothing process of the impulse response representing this context prior to its inversion has been applied. This technique allows reducing simultaneously the reverberation effect and the implementation complexity of the inverse filters [6].

In Chapter 1, we briefly describe a measurement method of acoustic impulse responses which is useful in the case where the anechoic chamber doesn't exist, and objective criterions calculating for room characterization.

In Chapter 2, we propose a modified version of the homomorphic method, to design stable minimum-phase inverse filters for partial equalization of non-minimum-phase acoustic impulse responses for small rooms. In this proposed approach, some of the dominant poles of the inverse filter transfer function are replaced by new ones before carrying out the inverse DFT. Listening tests results (MATLAB audio tools) obtained in the context of a car interior, show that with this technique, speech quality can be better controlled than when using standard method.

In Chapter 3, we propose a modification of the measured acoustic impulse response prior to its inversion. In large room, impulse responses present long duration. By applying an iterative simple complex smoothing, corresponding time reduced impulse responses are derived which conform to perceptual principles (preserving the initial transient data that are significant to the listener. That is, direct wave and first reflections). Successively, a smoothed impulse response is derived and then used to design a stable inverse filter for original impulse response equalization. Equalization results obtained in the context of an audio-conferencing room show the possibility to reduce the reverberation effect to a predetermined level, and a robustness indication when the listener changes position.

Chapter 4 presents some simulation results obtained when the smoothing process has been applied for a dereverberation system placed before a simple speech recognition system, and listening tests validation for a large audio-conferencing room, after implementation of reduced complexity of an equalization filter example on Texas Instruments development board.

Conclusions and future work are given in the end of this thesis.

**REFERENCES**

[1] A.Gilloire et J.P.Jullien, ‘L`acoustique des salles dans les télécommunications’, L`écho des RECHERCHES, No 127, 1er trimestre 1987.

[2] Neely S. T. and J. B. Allen, “Invertibility of a room impulse response”, J. Acoust. Soc. Am, vol. 66, no. 1, pp. 165-169, July 1979.

[3] B. D. Radlovic and R. A. Kennedy, “Non-minimum phase equalization and its subjective importance in room acoustics”, IEEE Transactions on Speech and Audio Processing, vol. 8, no. 6, Nov 2000.

[4] J. N. Mourjopoulos, "Digital Equalization of Room Acoustics", J. Audio. Eng. Soc, Vol 42, N°11, 1994 November.

[5] **M. Ahfir**, I. Kale, A. Krukowski and D. Berkani, “Partial Equalization of Non Minimum-Phase Impulse Responses”, EURASIP Journal on Applied Signal Processing, Volume 2006, n°13, ISSN 1110-8657.

[6] **M. Ahfir**, I. Kale, A. Krukowski and D. Berkani, “Room Equalization Based on Iterative Simple Complex Smoothing of Acoustic Impulse Responses”, 12° IEEE International Conference ICECS’05, TUNISIA, Vol 1, pp 256-259, Dec 2005.

Chapter 1

# ACOUSTIC IMPULSE RESPONSES MEASUREMENTS

## 1.1 INTRODUCTION

In room acoustics, impulse responses measurements should be around one second of duration, and 60 dB of magnitude. Classical techniques measurements which produce an impulse (balloon blowing up, pistol coup,…,etc) and record the impulse response do not allow to respond simultaneously to these requirements. Also there are considerable risks of distortions when impulses are applied to transducers. The measurement method described in this chapter and useful in the case where the anechoic chamber doesn't exist, overcomes these inconveniences by suppressing the impulse nature of the measurement test signal in one hand, and providing a substantial gain for the measured impulse response magnitude in the other hand. In this method, the measured impulse response is an identification method using cross correlation function. This identification method has been studied and compared to NLMS algorithm with decreasing step. The latest presents some advantages in the case of large rooms, but requires the use of the anechoic chamber if the test signal is not a maximum length sequence [1, 2]. Measured impulse responses can be analysed to calculate objective criterions for room characterisation.

## 1.2 MEASUREMENT METHOD

The test signal is a sequence called: 'pseudo-random binary sequence' (figure 1-1), and has the characteristics of a stationary white noise. The signal $x(t)$ applied to the input of the speaker results from a periodic pseudo-random sequence $x(n)$, which has its values in the interval $\{-1,1\}$, and generated by a shift register of $m$ stages. By choosing the number $m$, a 'maximum' and deterministic sequence periodic of $L = 2^m - 1$ can be generated. This sequence of $\pm 1$ is then converted to $\pm 1$ V volts at the sampling frequency $F_s$. The obtained signal $x(t)$ is periodic of $T = \frac{L}{F_s}$.

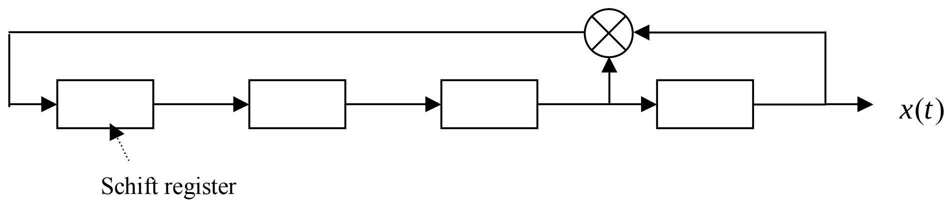



FIGURE 1.1: Maximum length sequence generated by a shift register of 4 stages

The test signal has a fundamental propriety of autocorrelation function. That is, the following periodic function:

$$R_{xx}(t) = \int_{-\infty}^{+\infty} x(t)x^*(t-t)dt = s_x^{\,2} d(t), \qquad (1.1)$$

$s_x^{\,2}$: variance of the test signal $x(t)$.

This propriety is used to identify the impulse response corresponding to a couple of speaker/microphone placed in the room. This system can be represented by convolution equation:

$$y(t) = x(t) * h(t) \quad , \qquad (1.2)$$

where $x(t)$ is the input, $y(t)$ is the output and $h(t)$ is the impulse response of the system.

The cross correlation function between $y(t)$ and $x(t)$ is written as:

$$R_{yx}(t) = \int_{-\infty}^{+\infty} y(t)x^*(t-t)dt \quad . \qquad (1.3)$$

The following relation results from the relations (1.2) and (1.3):

$$R_{yx}(t) = R_{xx}(t) * h(t), \qquad (1.4)$$

using the propriety (1.1), the relation (1.4) becomes:

$$R_{yx}(t) = s_x^{\,2} h(t). \qquad (1.5)$$

The output signal $y(t)$ of the microphone is recorded and sampled at the frequency $F_s$. The resulting discrete version of the recorded signal $y(n)$ has the same periodicity $L$ as the sequence $x(n)$. The sequence $h(n)$ can be then considered as the discrete version of the periodic impulse response $h(t)$.

The discrete function $h(n)$ is then obtained by cross correlation between the discrete function $y(n)$ and the discrete function $x(n)$. That is:

$$R_{yx}(k) = \frac{1}{pL}\sum_{n=1}^{pL} y(n)x^{*}(n-k), \tag{1.6}$$

where $R_{yx}(k)$ should be estimated for an integer number $p$ of periodicity $L$.

The discrete impulse response $h(k)$ will be identified by the following relation:

$$h(k) = \frac{R_{yx}(k)}{s_{x}^{2}}. \tag{1.7}$$

Cross correlation between the maximum length sequence applied to the speaker and the recorded signal by the microphone, represents then a measurement of the impulse response of the system. The measurement method can be summarised in the synoptic of Figure 1-2 [3].

Figures 1-3 and Figure 1-4 show respectively two examples of acoustic impulse responses measured in the car interior and in an audio-conferencing room.

## 1.3 OBJECTIVE CRITERIONS

To characterise the acoustical quality of any room, objective criterions are calculated from time energy distribution of the impulse response measured for an acoustic channel, where the couple source/receiver are fixed in the room. These objective criterions are related to perceptual criterions, where only subjective tests allow comparing their perceptual importance. There is a long and redundant list of objective criterions used by the acousticians [1]. But the most important in telecommunications are: reverberation time and clarity in time domain, and spectral deviation in frequency domain. In the following table, we show the results of these criterions calculated for the two previous examples of acoustic impulse responses.

The acoustical analysis of this table is of the psychoacoustics field. But these results can indicate that the reverberation phenomenon is more significant in a large room than in a small room. That is, great Reverberation Time and less Clarity.  This is can be seen in the Figure 1-3 and Figure 1-4 by very long impulse response duration.

**Table 1.1:** Objective criterions

| Objective criterions | Car interior | Audio-conferencing room |
|---|---|---|
| Reverberation Time (s) | 0.127 | 0.56 |
| Clarity 80 ms (dB) | 34.7 | 6.8 |
| Spectral Deviation (dB) | 5 | 8.31 |

For the equalization techniques evaluation, Spectral Deviation which is defined in the next chapter is the most used criterion for non-expert listeners to compare processed from unprocessed signals.

## 1.4 REFERENCES

[1] **M.Ahfir** et A. Benallal, ‘’Amélioration de la Méthode d’Identification dans la Mesure des Réponses Impulsionnelles Acoustiques ‘’, IEEA’97, International Conference, Vol Control Engineering, pp 60-63, University of BATNA, ALGERIA, Dec 1997.

[2] **M.Ahfir**, ‘’Modélisation et mesure des réponses impulsionnelles acoustiques’’, Mémoire de Magister, Université de BLIDA (ALGERIE), Octobre 1997.

[3] J.P.Jullien, A.Gilloire et A.Saliou, ‘’Mesure des réponses impulsionnelles en acoustique’’, Note technique NT/LAA/TSS/181, CNET – Lanion (France), Novembre 1984.

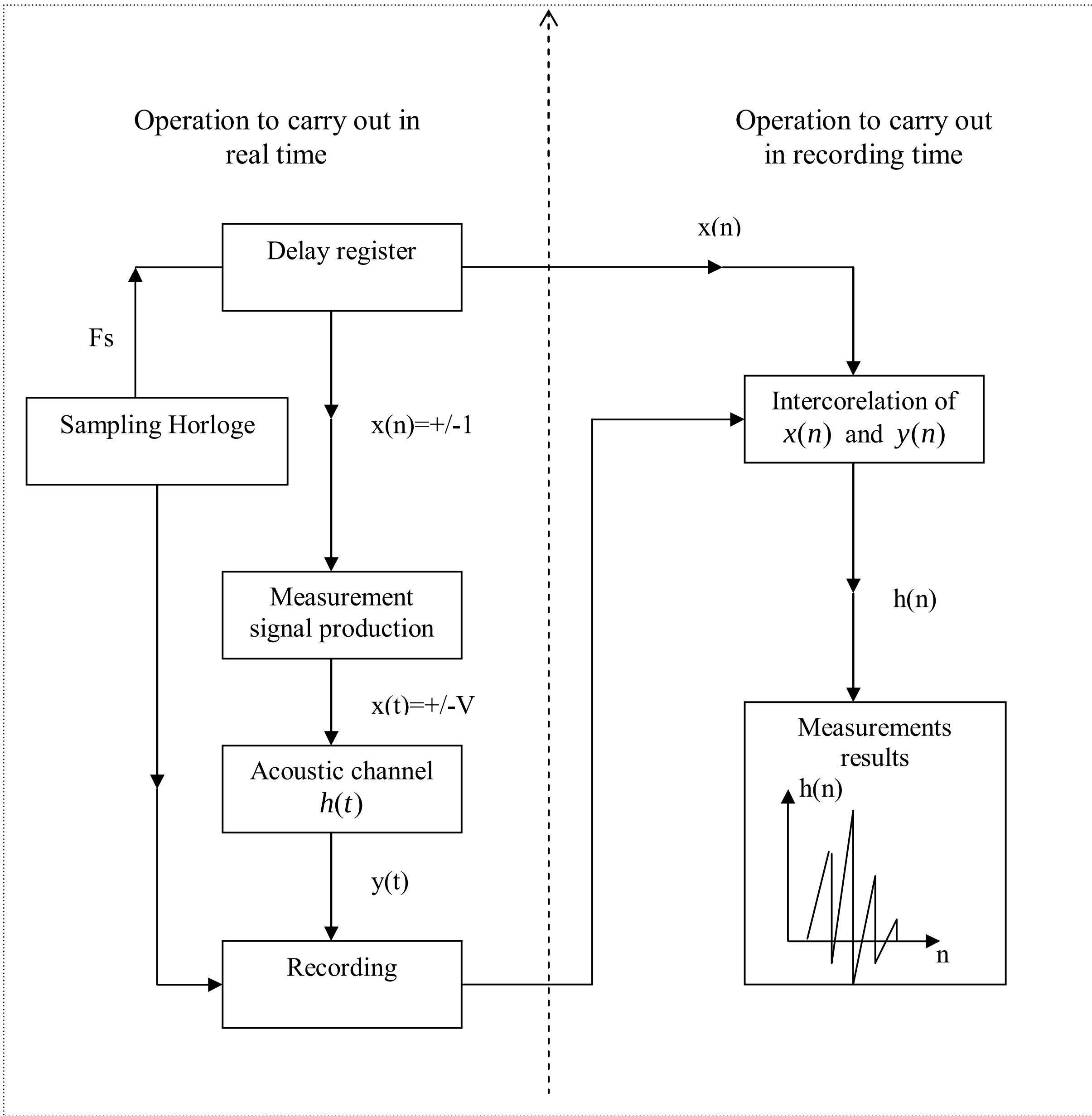



FIGURE 1.2: Synoptic of acoustic impulse responses measurements

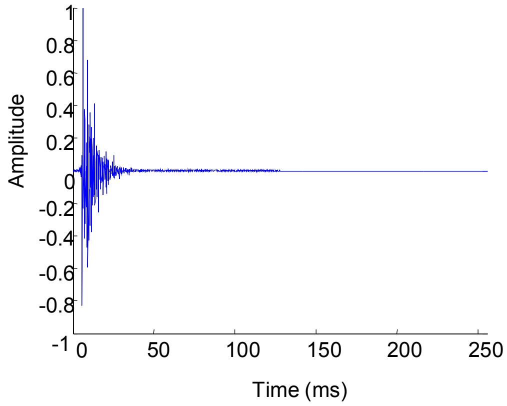



FIGURE 1.3: Acoustic impulse response measured in the car interior

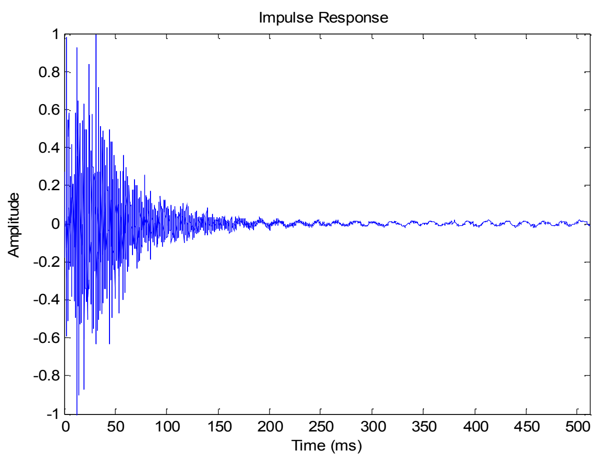



FIGURE 1.4: Acoustic impulse response measured in an audio-conferencing room

Chapter 2

# PARTIAL EQUALIZATION FOR SMALL ROOMS

## 2.1 INTRODUCTION

In sound-reproduction systems an equalization filter is often used to modify the frequency spectrum of the original source before feeding it to the loudspeaker. The purpose is to make the impulse response of the equalized sound-reproduction chain as close as possible to the desired one [1]. Single-point equalization based on a measured impulse response for a small room seems to be realistic because of a limited space for a listener placed inside. For example, at least for a driver placed inside the car interior. In principle direct equalization or direct inversion of mixed-phase (or non minimum-phase) measured impulse responses of the systems is not possible since it leads to unstable equalization filter realizations. Since any mixed-phase impulse response can be represented mathematically by the convolution of a minimum-phase and a maximum-phase (or all-pass) sequence [14], it is possible to derive and implement an approximate and stable inverse filter for such systems [10]. This is because a causal and stable sequence can invert the minimum-phase component of any mixed-phase sequence and an infinite acausal (anticipatory) and stable sequence can similarly invert the maximum-phase component of such sequences [10]. One method to design such a minimum-phase equalization filter is the homomorphic method based on the measured impulse response of the system. This method known as standard used for the case of single-point equalization is described in Section 2. In section 3, a modified version of the standard homomorphic method is proposed. It takes into account that the listener is able to detect gradual response variations of less than 0.5 dB [4, 5] and hence, is able to control the sound quality more accurately [16]. Section 4 shows the magnitude equalization performance results for an impulse response measured in a car interior using both objective and subjective measurements.

## 2.2 STANDARD HOMOMORPHIC METHOD

A non minimum-phase discrete impulse response, *h(n)*, of a system can be described as [14]

$$h(n) = h_{mp}(n) \otimes h_{ap}(n)\,, \tag{2.1}$$

where ⊗ denotes the discrete convolution. This can be shown in the frequency domain as

$$H(k) = H_{mp}(k) H_{ap}(k)\,, \tag{2.2}$$

where $h_{mp}(n)$ is a minimum-phase sequence, such that its DFT, $H_{mp}(k)$, satisfies the relation

$$\left|H_{mp}(k)\right| = \left|H(k)\right|, \tag{2.3}$$

where *H(k)* is the DFT of *h(n)* given by

$$H(k) = \sum_{n=0}^{N-1} h(n) e^{-j\frac{2p\,kn}{N}}\,, \tag{2.4}$$

where *N* is the length of *h(n)* and $h_{ap}(n)$ is an all-pass sequence of $|H_{ap}(k)|=1$, for $k=0,1,\ldots,N-1$.

The convolution operation of $h_{mp}(n)$ and $h_{ap}(n)$ can be expressed as the algebraic addition of their corresponding complex cepstra $\hat{h}_{mp}(n)$ and $\hat{h}_{ap}(n)$ by the homomorphic transformation [2]. This leads to a decomposition of a non minimum-phase impulse response into its minimum-phase and all-pass components. The standard homomorphic method algorithm is outlined as follows [3, 4 and 15]:

1. Compute the DFT of *h(n)*.
2. Compute

$$\hat{H}(k) = \log\left|H(k)\right|. \tag{2.5}$$

3. Compute the real part of the complex cepstrum of *h(n)*,

$$\hat{h}(n) = \frac{1}{N}\sum_{k=0}^{N-1} \log\left|H(k)\right| e^{j\frac{2p\,kn}{N}}\,, \tag{2.6}$$

for $n=0, 1\ldots,N-1$.

4. Compute the corresponding real cepstrum of the minimum-phase $h_{mp}(n)$,

$$\hat{h}_{mp}(n) = \begin{cases} \frac{\hat{h}(n)}{L} & n = 0, \frac{N}{2} \\ \frac{2\hat{h}(n)}{L} & 1 \le n \le \frac{N}{2} \\ 0 & \frac{N}{2} < n \le N-1 \end{cases}, \qquad (2.7)$$

where $L$ is a positive real parameter [15].

5. Compute the DFT of $\hat{h}_{mp}(n)$,

$$\hat{H}_{mp}(k) = \sum_{n=0}^{N-1} \hat{h}_{mp}(n) e^{-j\frac{2p\,kn}{N}} . \qquad (2.8)$$

6. Compute *the minimum-phase part* $H_{mp}(k)$,

$$H_{mp}(k) = \exp(\hat{H}_{mp}(k)) . \qquad (2.9)$$

7. Compute the equalized response, $H_{eq}(k)$,

$$H_{eq}(k) = H(k)G_{mp}(k) , \qquad (2.10)$$

where $G_{mp}(k)$ represents the inverse of $H_{mp}(k)$,

$$G_{mp}(k) = \frac{1}{H_{mp}(k)} . \qquad (2.11)$$

In the time domain, this is equivalent to a deconvolution,

$$h_{eq}(n) = h(n) \otimes g_{mp}(n) , \qquad (2.12)$$

with $g_{mp}(n)$, being the inverse DFT of $G_{mp}(k)$.

In the case of $L=1$, the algorithm corresponds to a magnitude equalization. If a sufficiently large number $N$ is used for DFT computation, the effect of magnitude distortion caused by the system can be perfectly removed in practice by convolving $h(n)$ with the inverse minimum-phase impulse response $g_{mp}(n)$ [3, 10].

To avoid the impulse response length of $g_{mp}(n)$ to be truncated given a fixed number $N$ of DFT computation, an iterative version of the standard method can be used with an optimum integer number $L$ of iterations. This allows reduction in time domain aliasing, when convolving $h(n)$ with $g_{mp}(n)$ [ 15]. This version can be summarized as follows:

$$h(n)^{i+1} = h(n)^{i} \otimes gmp^{\frac{1}{L}}(n),$$

where $i = 0,...,L-1$ , $h^0(n) - h(n)$ *and* $h^L(n) = hap(n)$ . $gmp^{\frac{1}{L}}(n)$ being the partial minimum-phase inverse impulse response.

The effect of phase distortion can also be solved by convolving the all-pass sequence, $h_{ap}(n)$ (obtained after deconvolution of $h(n)$ with $g_{mp}(n)$) with its time reversed version, $h_{ap}(-n)$, [4, 9]. As a result, implementation of such combined equalization (complete equalization) requires very long FIR inverse filters. But this is not always required in practice. For this reason, the equalization of the all-pass component (phase equalization) will not be considered in this work.

In the case of $L>1$, the algorithm corresponds to a partial magnitude equalization. This requires a shorter FIR filter to keep the phase distortion below the threshold of audibility [4].

Sometimes, the frequency response of the system, $H(k)$, and hence its minimum-phase part, $H_{mp}(k)$, can be represented by a low number of isolated dominant zeros. In such case increasing the parameter $L$ by a significant value during the control process may shorten the length of the equalization filter too, resulting in an unsatisfactory equalization performance. This is because an increase in $L$ results in a decrease of all the radii of the complex poles of $G_{mp}(k)$ together according to the relation derived from (3), (7) and (9) [Appendix A] ,

$$\log\left|G_{mp}(k)\right| = -\frac{1}{L}\log\left|H(k)\right| . \tag{2.13}$$

This means that the complex poles of $G_{mp}(k)$ appear to be pushed together towards the origin of the unit circle.

In the next section, we propose an alternative approach in which, instead of pushing all the poles of $G_{mp}(k)$, we push the most dominant of them selectively and slightly towards the origin of the unit circle by decreasing the corresponding high values of the $Q$ factors (values of the steady-state resonances). This allows controlling the magnitude equalization performance more precisely – especially applicable in practice. The reason is that the listener is able to detect gradual response variations of less than 0.5 dB [4, 5]. Furthermore, the proposed technique is advantageous when the parameter $L$ cannot be calculated theoretically, e.g. for the case of the direct inverse filtering (no cepstral analysis, that is, no steps 2 to 6) of a small reverberant room where the dominant poles can be identified even if they are closely spaced [8, 11].

## 2.3 MODIFIED VERSION

A replacing method of some dominant poles of the inverse minimum-phase function $G_{mp}(k)$ is described in this section. They are identified using the standard method (*L=1*) and then replaced before doing the inverse DFT in order to calculate the corresponding discrete time sequence $\tilde{g}_{mp}(n)$ representing the impulse response of the new equalization filter.

The z transform function of a complex pole pair is expressed as [8, 12]

$$H_p(z) = \frac{1}{\left(1-|a|e^{jq}z^{-1}\right)\left(1-|a|e^{-jq}z^{-1}\right)},$$

or (2.14)

$$H_p(z) = \frac{1}{1-2|a|\cos q\, z^{-1}+|a|^2 z^{-2}},$$

where $|a|$ is the pole radius in the z plane and $2p\dfrac{f_p}{f_s}$ is its phase angle with $f_s$ being the sampling frequency and $f_p$ the frequency of the complex pole.

Taking the inverse z transform of $H_p(z)$ the corresponding impulse response is [8, 12]

$$h_p(n) = \frac{|a|^n \sin(nq+q)}{\sin(q)}u(n), \tag{2.15}$$

where *u(n)* is a unit step function.

The transfer function of the selective filter for a complex pole pair is

$$H_s(z) = \frac{\tilde{H}_p(z)}{H_p(z)}, \tag{2.16}$$

where the transfer function $\tilde{H}_p(z)$ contains a new complex pole pair at the same frequency of the old pair but at a desired smaller radius, $|\tilde{a}|$. This technique allows us to decrease selectively the $Q$ factors values of a low order of isolated pole pairs in the frequency response $G_{mp}(k)$. The new inverse minimum-phase function becomes

$$\widetilde{G}_{mp}(z) = G_{mp}(z) H_s^{(1)}(z) \ldots H_s^{(P)}(z) \ , \tag{2.17}$$

and its discrete version

$$\widetilde{G}_{mp}(k) = G_{mp}(k) H_s^{(1)}(k) \ldots H_s^{(P)}(k) \, , \tag{2.18}$$

where $P$ is the number of identified and replaced dominant pole pairs from $G_{mp}(k)$, and $H_s^{(p)}(k)$, $p=1,\ldots,P$ are the sampled frequency responses of selective filters equal to the number of replaced pole pairs, $P$.

This function is then inverted using the inverse DFT in order to obtain its discrete time domain equivalent $\widetilde{g}_{mp}(n)$, shorter from $g_{mp}(n)$ calculated by standard method for $L=1$ and longer than $g_{mp}(n)$ obtained for $L>1$.

One method to identify frequencies of the isolated poles is to iteratively search for the increased magnitude response level of $G_{mp}(k)$ caused by poles (peaks) residing within the frequency range of interest (in our case below $4\ kHz$). Successively, a maximum magnitude level $G_{mp}(f_p)$ corresponding to the highest pole frequency $f_p$ is found. This technique was found robust even in the case of very closely spaced poles [8, 11]. After determining the frequency $f_p$ of the highest pole, the corresponding pole radius must be determined based on the $Q$ factor value according to the following relation, since our work here is restricted to a low order of isolated poles [6, 7, 8, and 13],

$$Q = G_{mp}(f_p) = \frac{1}{1-|a|} \ . \tag{2.19}$$

The replacing method means that the dominant poles of $G_{mp}(k)$ are identified one by one and then replaced iteratively by new ones, where each corresponds to a desired $\widetilde{Q}$ factor, $\widetilde{Q} = \frac{1}{1-|\widetilde{a}|}$, starting from the most dominant one.

The implementation algorithm of the proposed modified method (useful for partial magnitude equalization), is as follows:

1. Compute the steps 1 to 6 as in the standard method for $L=1$.
2. Compute the inverse minimum-phase

$$G_{mp}(k) = \frac{1}{H_{mp}(k)}\,, \tag{2.20}$$

using $G_{mp}(k) = \widetilde{G}_{mp}(k)$.

3. Set *p to 1*.
4. Estimate the most dominant pole from $\widetilde{G}_{mp}(k)$ as described above, (determine $f_p$ and $|a|$).
5. Design its selective filter using (3.16).
6. Replace the estimated pole from $\widetilde{G}_{mp}(k)$ using (3.18).
7. Increment *p=p+1* and repeat the steps 4, 5 and 6 until *p=P*.
8. Compute $\widetilde{g}_{mp}(n)$ as inverse DFT of $\widetilde{G}_{mp}(k)$.
9. Compute the equalized response $H_{eq}(k)$,

$$H_{eq}(k) = H(k)\widetilde{G}_{mp}(k)\,. \tag{2.21}$$

In the time domain, this is equivalent to a deconvolution

$$h_{eq}(n) = h(n) \otimes \widetilde{g}_{mp}(n)\,. \tag{2.22}$$

In the next section we present the performance evaluation of the magnitude equalization performed by the proposed version as compared to that from standard method, using both objective measures based on an error criterion and subjective tests of speech quality.

## 2.4 RESULTS

In order to assess the performance of our algorithms, we used a frequency domain error criterion, which estimates the standard deviation of the magnitude response from a constant level [4]. The error criterion *D(dB)* is given as follows

$$\Delta = \left[\frac{1}{N}\sum_{k=0}^{N-1}\left(10\log_{10}\left|H_{eq}(k)\right| - H_m\right)^2\right]^{1/2}, \tag{2.23}$$

where

$$H_m = \frac{1}{N}\sum_{k=0}^{N-1} 10\log_{10}\left|H_{eq}(k)\right|. \tag{2.24}$$

Two examples of non minimum-phase impulse responses were used to compare both algorithms. The first one was synthetic, used just to enlighten the replacing approach of a low order of isolated poles, and the second one used real measurements taken in a car interior. We also introduced in the proposed version a real parameter $l\,(l>1)$, in order to selectively decrease the highest $Q$ factors of dominant poles. This means that the new replacement poles correspond to desired $\tilde{Q}$ factors, $\tilde{Q}=Q/l$ .

**2.4.1 Synthetic impulse response**

We first considered a simple synthetic impulse response. This was obtained by successive convolution of six known zeros sequences somewhat isolated in the complex z plane (Figure 3-1) and with at least one placed outside the unit circle, which make the impulse response a non minimum-phase one. These are defined as ( $f_S=8kHz$ ):

$|a|=0.99$ at $f_p=200Hz$

$|a|=0.99$ at $f_p=1000Hz$

$|a|=0.85$ at $f_p=1500Hz$

$|a|=0.70$ at $f_p=2000Hz$

$|a|=1.5$ at $f_p=2500Hz$

$|a|=0.95$ at $f_p=3000Hz$

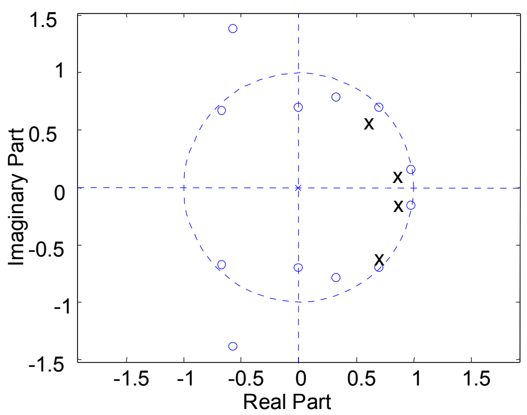



FIGURE 2.1: Complex z plane of six known zeros

(Poles after inversion of the minimum-phase part):

(o) Poles pairs and (x) replacing poles pairs.

Figure 2-2 shows the inverse frequency response $G_{mp}(k)$ from which the two *(P=2)* most dominant poles are estimated and corresponding to:

$$|a_1|=0.9947 \quad at \quad f_1=200Hz$$

$$|a_2|=0.9915 \quad at \quad f_2=1000Hz$$

These two *(P=2)* poles are selectively replaced by two new poles of smaller radius corresponding to $Q_1/l$ and $Q_2/l$ factors respectively, with a significant value of $l$, ($l=2$) (Figure 3-1). In Figure 3-2, even with some error in the estimation of poles, we still can observe a decrease in the $Q$ factors depending on the position of the new poles. This corresponds to a reduction of $\tilde{g}_{mp}(n)$ length when compared to $g_{mp}(n)$. When using the standard method and considering the same significant value of $L$ ($L=l=2$), we can see in Figure 2 that all the poles have been pushed together towards the origin of the unit circle too, resulting also in the reduction of the $g_{mp}(n)$ length, but this is considerably shorter than that of $\tilde{g}_{mp}(n)$.

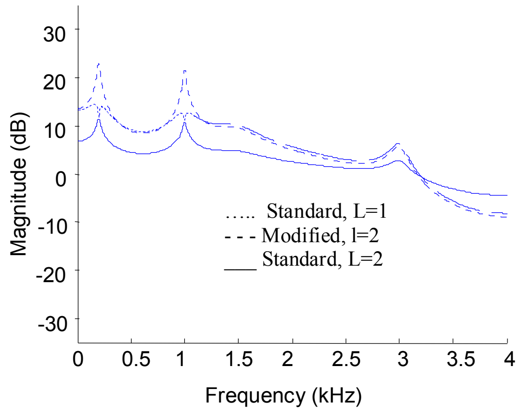



FIGURE 2.2: Inverse frequency response $G_{mp}(k)$ calculated using different methods.

The evaluation of the objective error criterion for this example is not considered because of its little practical interest.

### 2.4.2 Practical impulse response

A real impulse response was measured for the car interior at a sampling frequency of 8 kHz. A record of 1024 samples was zero padded up to $N=2048$. This impulse response is shown in Figure 2-3. In Figure 2-4 an unstable direct inverse impulse response is shown, demonstrating its non minimum-phase character.

Figure 5 shows the inverse minimum-phase frequency response $G_{mp}(k)$. It was calculated using the standard method ($L=1$). The most dominant pole can be clearly seen there. The search was limited to a single pole, i.e. $P=1$, such that $|a_1|=0.9993$ at $f_1=70.38\ Hz$ that caused the inverse minimum-phase impulse response $g_{mp}(n)$ to be of a very long duration (Figure 2-7).

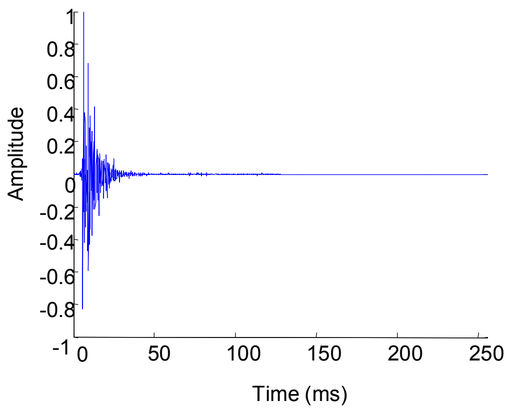



FIGURE 2.3: Impulse response measured in the car interior.

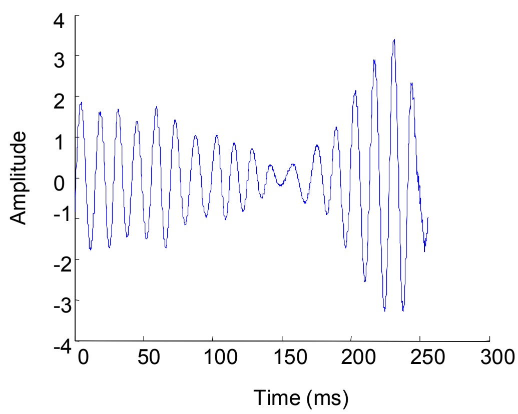



FIGURE 2.4: Direct inverse impulse response.

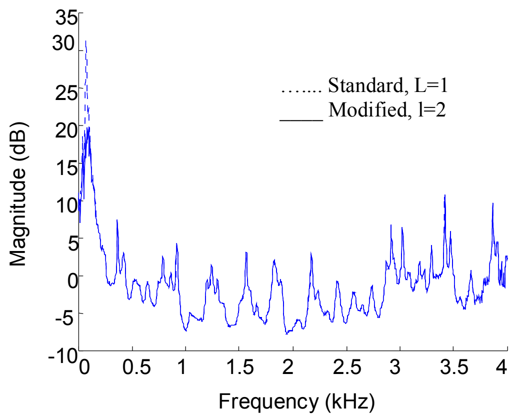



FIGURE 2.5: Inverse minimum-phase frequency response $G_{mp}(k)$ calculated by the different methods.

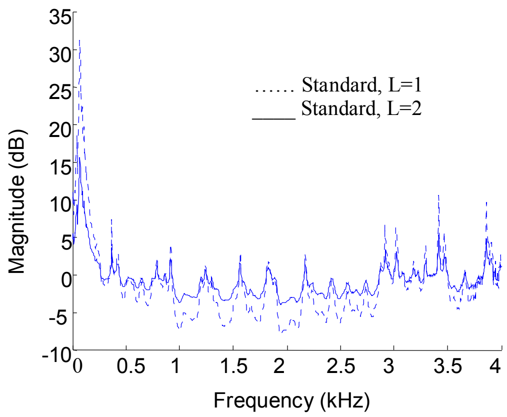



FIGURE 2.6: Inverse minimum-phase frequency response $G_{mp}(k)$ calculated by the different methods.

When using the standard version for a significant value of $L$, *(L=2)*, (Figure 3-6) we observed that all the poles of $G_{mp}(k)$ were pushed together towards the origin of the unit circle too, resulting in an inverse minimum-phase impulse response $g_{mp}(n)$ (Figure 3-7) to be reduced in time too.

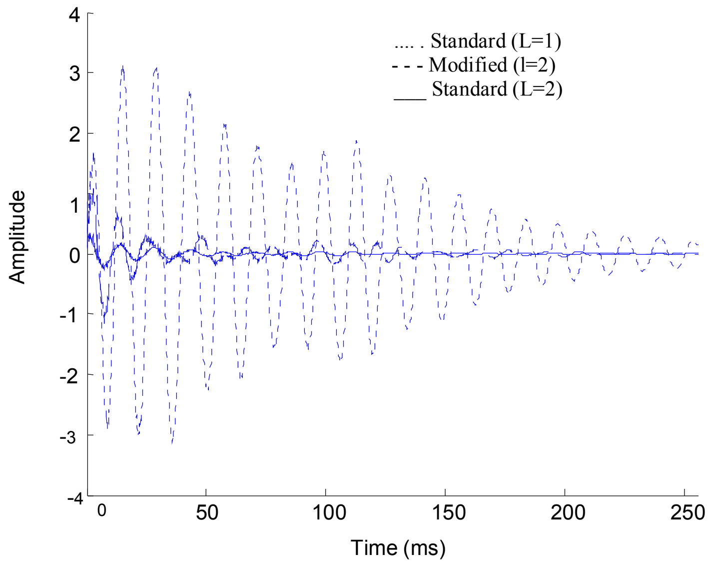



FIGURE 2.7: Inverse minimum-phase impulse responses $g_{mp}(n)$ calculated by the different methods.

When using a modified version, in order to gradually reduce the length of the inverse minimum-phase impulse response $g_{mp}(n)$, only the most dominant pole needed to be replaced by a new pole with smaller radius. This pole corresponded to a $Q/l$ factor with the same value of $l$, ($l = L = 2$), but at the same frequency. In Figure 2-5 we can see the inverse minimum-phase frequency response of $\widetilde{G}_{mp}(k)$, where only the most dominant pole appears to be pushed towards the origin, with $l = 2$.

Figure 2-7 also shows the corresponding inverse minimum-phase impulse response of $\widetilde{g}_{mp}(n)$. Interestingly, its duration is not reduced here too. This may correspond to a desired magnitude equalization (Figure 2-8), if the system impulse response was

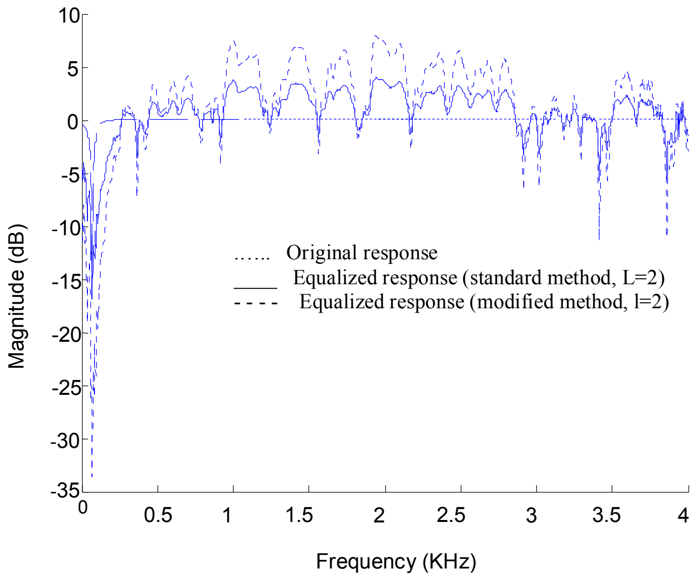



FIGURE 2.8: Magnitude response equalization

minimum-phase (no phase distortion effects). This is because the magnitude spectrum of the second case (modified method, $l = 2$, $\Delta(dB) = 0.7$) is flatter than that of the first case (standard method, $L = 2$, $\Delta(dB) = 2.4$). This means less magnitude distortion of the system.

### 2.4.3 Performance testing

The experiment was performed by developing models in Matlab and Simulink and carrying out listening tests using a headphone ((Logitech – Premium Stereo Headset). A reproduced speech signal of few second in duration was generated by filtering a clean speech (male and female measured in anechoic chamber) by the measured impulse response of the car interior (Figure 2-3). In order to avoid undesirable convolution effects, we considered a sufficient large number *N=8192* for DFT computations. The reproduced speech signal was then filtered using equalizing filters calculated by the standard method with *L=1* and *L=2* (Figure 2-7) and the modified version $(P=1)$ respectively. For the latter case, the inverse impulse responses corresponded to each error criterion, function of the parameter $l$ such as that of Figure 2-7 with $l=2$ for example. Test signals were played to ten untrained listeners with normal hearing at a comfortable listening level. The qualitative assessment of the test signals was based on subjective judgment of three listening sessions per each recording scheduled on six consecutive days.

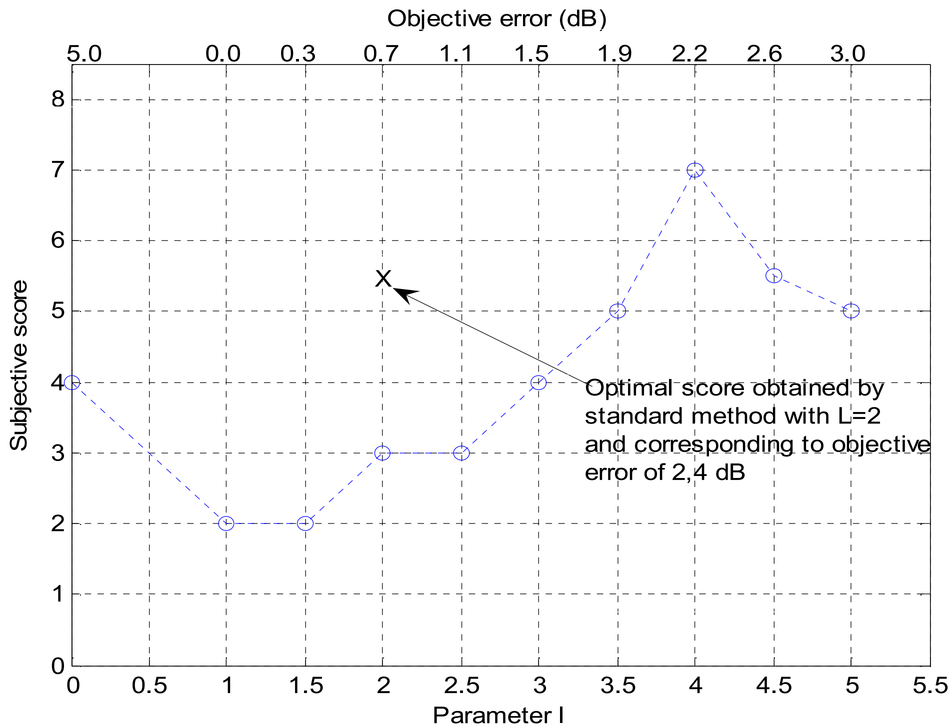



FIGURE 2.9: Subjective scores of the sound quality as a function of the parameter $l$.

Each circle represents the average of 180 observations: 18 for each of 10 listene

The first signal was always chosen to be clean speech, while the reproduced unequalized and partially equalized speech signals were played in random order. The reproduced speech signals corresponded to the objective error criterions of 5 dB (unequalized signal for $l = 0$), 0 dB (magnitude equalized signal for $l = 1$) and $0.3 \le \Delta(dB) \le 3$ (partially equalized signals for $l > 1$) respectively. The quantification of subjective judgments was performed according to the following scale [4]:

7, 8 – Good

5, 6 – Fair

3, 4 – Poor

1, 2 – Bad

Number 8 denotes a sound quality equivalent to the clean speech. The final result was calculated as a mean of the individual listening results (18 each) for each of 10 subjects. The results are shown in Figure 3-9 as a function of parameter $l$, ranging from $l = 0$ (unequalized signal) to $l = 5$ (partially equalized signal).

The results confirmed those reported in [4] with higher accuracy. The highest score corresponds to the optimal quality of speech. That means no perception of phase distortion (like a bell chime sounded at the background, when $l < 3$), no echo and less magnitude distortion caused by the system.

The results also show the sensitivity of the listener's ear to small gradual response variations (a variation of less than 0.5 dB of objective error corresponds to a significant variation of subjective score of the sound quality); although the participants in the experiment were non-expert listeners.

## 2.5 CONCLUSIONS

In this chapter a modified version of the standard homomorphic method for small rooms equalization is presented. This version is useful in case of partial magnitude equalization, where the dominant zeros density of the system is not very high. Although it is used in this work as an additional optimizing tool for the psycho acoustic quality measurement of speech, this alternative approach is advantageous in case of the direct inverse filtering (minimum-phase system) when perfect equalization of a small reverberant room is not desired.

## 2.6 REFERENCES


[1] Elliott S. J. and P. A. Nelson, "Multiple-point equalization in a room using adaptive digital Filters", J. Audio Eng. Soc., vol. 37, no. 11, Nov 1989.

[2] Oppenheim A. V. and R. W. Schafer, *Discrete Time Signal Processing*, New Jersey: Prentice Hall, 1989.

[3] Neely S. T. and J. B. Allen, "Invertibility of a room impulse response", J. Acoust. Soc. Am, vol. 66, no. 1, pp. 165-169, July 1979.

[4] B. D. Radlovic and R. A. Kennedy, "Non-minimum phase equalization and its subjective importance in room acoustics", IEEE Transactions on Speech and Audio Processing, vol. 8, no. 6, Nov 2000.

[5] Fielder L. D., "Analysis of traditional and reverberation-reducing methods of room equalization", J. Audio. Eng. Soc., vol. 51, no. 1/2, Jan/Feb 2003.

[6] Haneda Y., S. Makino and Y. Kaneda, "Common acoustical pole and zero modelling of room transfer functions", IEEE Transactions on Speech and Audio Processing, vol 2, no. 2, April 1994.

[7] Kunt M., "Traitement numérique des signaux", Dunod, Paris, 1981.

[8] Makivirta A., P. Antsalo, M. Karjalainen and V. Valimaki, "Modal equalization of loudspeaker-room responses at low frequencies", J. Audio. Eng. Soc, vol. 51, no. 5, May 2003.

[9] Preis D., "Phase Distortion and Phase Equalization in Audio Signal Processing–A Tutorial Review", J. Audio Eng. Soc., vol. 30, no. 11, Nov 1982.

[10] J. N. Mourjopoulos, "Digital Equalization of Room Acoustics", J. Audio. Eng. Soc, Vol 42, N°11, 1994 November.

[11] Karjalainen M., P. A. A. Esquef, P. Antsalo, A. Makivirta and V. Valimaki, "Frequency-zooming ARMA modelling of resonant and reverberant systems", J. Audio Eng. Soc., vol. 50, no. 12, Dec 2002.

[12] Bellanger M., "Traitement numérique du signal", Dunod, Paris, 1998.

[13] J. R. Hopgood and P. J. W. Rayner; "Blind Single Channel Deconvolution Using Nonstationary Signal Processing", IEEE Transactions on Speech and Audio Processing, vol 11. pp 476-488, March 2003.

[14] A. V. Oppenheim and R. W. Schafer, "Digital Signal Processing " (Prentice-Hall, Englewood Cliffs, NJ, 1975).

[15] B. D. Radlovic and R. A. Kennedy, “Iterative cepstrum-based approach for speech dereverberation ", ISSPA` 99, Brisbane, Australia, 22-25 August 1999.

[16] **M. Ahfir**, I. Kale, A. Krukowski and D. Berkani, “Partial Equalization of Non Minimum-Phase Impulse Responses”, EURASIP Journal on Applied Signal Processing, Volume 2006, n°13, ISSN 1110-8657.

Chapter 3

# PARTIAL DEREVERBERATION FOR LARGE ROOMS

## 3.1 INTRODUCTION

Equalization for large rooms is dereverberation. This is because reverberation phenomenon is more significant in such rooms. This means, very long acoustic impulse responses duration. Direct equalization based on inverse filtering of measured mixed-phase impulse responses introduces a number of theoretical and practical problems (Chapter3 - ideal digital equalization – perfect dereverberation). Particularly, the required extremely long lengths of the inverse filters (magnitude and phase equalization) and their sensitivity to possible changes in the listener position inside the room [1, 2]. When the smoothed room impulse response was used for the design of `inverse filters` for audio or acoustic digital equalization applications, it was found that the effect of smoothing was desirable since it was allowing the design of filters with lower sensitivity – from the perceptual point of view - to possible changes in the listener position inside the room [3, 4, 5]. The work presented here is based on the concept of complex smoothing performed by the simple case of constant bandwidth filters for `inverse filter` design for acoustic impulse responses equalization. This concept is adopted because modified impulse response functions could be derived, which would present functions of reduced complexity and also be in agreement with the perceptual principles. The work presented in this chapter is organized as follows: Section 2 presents the theory of the simple complex smoothing [3]. In Section 3 an equalization filter design method is proposed, which is based on iterative simple complex smoothing [9]. Section 4 shows the magnitude equalization performance results achieved using this proposed method in the context of a real audio-conferencing room (cnet-LANNION-FRANCE).

### 3.2 SIMPLE COMPLEX SMOOTHING

Let us consider a discrete-time room impulse response $h(n)$ and its discrete-frequency response $H(k)$, where $0 \le n \le N-1$, $0 \le k \le N-1$ and $N$ is the number of samples, representing the impulse response length. Then the complex smoothing operation for the simple case of constant bandwidth filter can be described as a circular convolution [3]:

$$H_{cs}(k) = H(k) \otimes W_{sm}(k) \tag{3.1}$$

Where $\otimes$ denotes the operation of circular convolution and $W_{sm}(k)$ is a spectral smoothing function having the general form of a low-pass filter.

This is equivalent in the time domain to:

$$h_{cs}(n) = h(n) N w_{sm}(n) \tag{3.2}$$

where $w_{sm}(n)$ being the inverse DFT sequence of $W_{sm}(k)$.

To consider this spectral smoothing function in a half-window sense for both parts of the symmetric spectrum $W_{sm}(k)$ must be written as:

$$W_{sm}(k) = \begin{cases} \dfrac{b-(b-1)\cos(pk/m)}{2b(m+1)-1}, & k = 0,1,\ldots,m \\ \dfrac{b-(b-1)\cos(p(k-N)/m)}{2b(m+1)-1}, & k = N-m, N-(m-1),\ldots, N-1 \\ 0 & k = m+1,\ldots, N-(m+1) \end{cases} \tag{3.3}$$

Where $m$ (samples) is defined as the smoothing index corresponding to the length of the half window. When $b = 1$, this function represents an ideal low-pass filter (rectangular frequency smoothing function),

$$W_{sm}(k) = \begin{cases} \dfrac{1}{2m+1}, & k \in \{0,1,\ldots,m\} \cup \{N-m, N-(m-1),\ldots, N-1\} \\ 0 & k \in \{m+1,\ldots, N-(m+1)\} \end{cases} \tag{3.4}$$

In this case the corresponding time window function which is represented with a *sinc* function $w_{sm}(n)$ can be evaluated easily as:

$$w_{sm}(n) = \frac{1}{N} \frac{\sin(pn(2m+1)/N)}{pn(2m+1)/N} \tag{3.5}$$

Given that $H(k)$ is a complex function; in general $W_{sm}(k)$ should be a complex function. However, the preceding expressions represent it as a real function, assuming it to be a zero-phase function. This assumption was adopted because of physical considerations, since with smoothing it is required to avoid imposition of any unwanted effects on the phase of the original function and it is also desirable to maintain the half-window time profile appropriate for capturing transient data of the acoustic impulse response that are significant to the listener (direct signal and first reflections).

### 3.3 EQUALIZATION FILTER DESIGN BASED ON ITERATIVE SIMPLE COMPLEX SMOOTHING

In order to overcome the known problems of direct equalization as stated in the introduction, we propose in this section an equalization method based on iterative simple complex smoothing of room impulse response. In this method the original room impulse response $h(n)$ is assumed to be the result of smoothing operation in the time domain of an initial one (its length was $2N$ ) by a half smoothing window $w_{sm}(n)$ of $N$ samples. Then in the first iteration $h(n)$ will be halved in length after the application of a half smoothing window $w_{sm}(n)$ of $\frac{N}{2}$ samples. As a result a reduced-length complex smoothed impulse response $h_{cs}(n)$ will be produced and a minimum-phase inverse impulse response $g_{cs,\,mp}(n)$ for smoothed response magnitude equalization will be evaluated by using the homomorphic method [6], since the smoothed response may present a mixed-phase character. This minimum-phase inverse impulse response $g_{cs,\,mp}(n)$ will be the `equalization filter`, employed on the original acoustic impulse response $h(n)$ according to the following expression:

$$h_{eq}(n) = h(n) \otimes g_{cs,\,mp}(n) \tag{3.6}$$

The iteration process is repeated only if the result of the smoothing operation will be meaningful, that is, the early time components (direct signal and first reflections) of the already smoothed impulse response $h_{cs}(n)$ must be preserved when halving it again with a half smoothing window $w_{sm}(n)$. The procedure may be stated as follows:

1. Set the iteration index $i$ to 0 and choose the parameters $b$ and $m$ .
2. Compute the spectral smoothing window $W_{sm}(k)$ using (3) for $k = 0,1,2,....,\frac{N}{2^i}$ .

3. Compute the corresponding time smoothing window $w_{sm}(n)$ by using the IDFT and apply it for $n = 0,1,2,....,\frac{N}{2^{i+1}}$.

4. Compute the smoothed impulse response $h_{sm}^{i+1}(n)$ as:

$$h_{cs}^{i+1}(n) = h_{cs}^{i}(n)\frac{N}{2^{i}} w_{sm}(n) \tag{3.7}$$

where $h_{cs}^{0}(n) = h(n)$.

5. Evaluate the corresponding minimum-phase inverse impulse response $g_{cs,\,mp}^{i+1}(n)$ using the homomorphic method.

6. Compute the equalized room impulse response $h_{eq}(n)$ as:

$$h_{eq}(n) = h(n) \otimes g_{cs,\,mp}^{i+1}(n). \tag{3.8}$$

7. Increment $i$, $i = i + 1$ and go to step 2.

### 3.4 RESULTS

The proposed method was applied to a measured acoustic impulse response of an audio-conferencing room (typical of a large room) at a sampling frequency of 16 kHz and recorded to $N = 8192$ (Figure 3-1) [8]. In this example three iterations ($i = 0,1,2$) of the smoothing process using an initial half smoothing window (Figure 3-2- $b = 0.5$ and $m = 3$) was needed to preserve the initial transient portion of the acoustic impulse response. A reduced-length of the smoothed response $h_{cs}^{3}(n)$ (Figure 3- 3) was produced and then used to construct the appropriate minimum-phase inverse impulse response $g_{cs,\,mp}^{3}(n)$ by using an iterative version of the homomorphic method [7] (Chapter 2) with an optimum number of iterations $L = 8$. This means that $g_{cs,\,mp}^{3}(n)$ is constructed by 8 successive convolutions of partially calculated minimum-phase inverse impulse responses $g_{cs,\,mp}^{1/8}(n)$ (Figure 3-4). The particular choice of this iterative version over the standard method is because of reduced time domain aliasing given a fixed number $N$ for DFT computation.

**Objective and subjective evaluations**

In order to assess the performance of our algorithm, we used an objective frequency domain error criterion, which estimates the standard deviation of the magnitude response from a constant level [2]. The error criterion $\Delta(dB)$ is given as follows:

$$\Delta = \left[ \frac{1}{N} \sum_{k=0}^{N-1} \left( 10 \log_{10} |H(k)| - H_m \right)^2 \right]^{1/2} \tag{3.9}$$

where:

$$H_m = \frac{1}{N} \sum_{k=0}^{N-1} 10 \log_{10} |H(k)| \tag{3.10}$$

and evaluated for both the original and equalized responses.

But first, let us consider the case of single-point equalization, where the objective error criterion was calculated function of the number of iteration of the smoothing process, $i$ (Table 4-1). Corresponding inverse impulse responses of the equalization filters which have been successively produced are shown in the Figure 3-5, where we can see a successive reduction in complexity function of $i$.

Theses results indicate that reverberation phenomenon is significantly removed after the first ($i = 0$) and second ($i = 1$) iteration, and significantly reduced after the third ($i = 2$) iteration as compared to that of the original response.

**Table 3.1:** Objective error criterion (dB) evaluated for both original and equalized responses function of $i$ (Single-point equalization)

| **Number of iteration** $i$ | $\Delta(dB)$ |
|---|---|
| 0 | 1.8 |
| 1 | 2.86 |
| 2 | 3.54 |
| **Original response** | 8.31 |

Listening tests have been carried out for subjective evaluation. We used a headphone (Logitech – Premium Stereo Headset), Matlab's Audio Tools and a reverberated anechoic speech (obtained by convolution of anechoic speech with the acoustic impulse response of Figure 3-1) as original test signal. The original test signal was convolved with the impulse responses of the equalization filters corresponding to first, second and third iteration respectively (Figure 3-5). Listening test processed signals were produced which respectively correspond to first, second and third iteration of the smoothing process. Subjective results indicated:

**- For the 1er case ($i = 0$):**

The effect of magnitude distortion was perceived as significantly removed from the reverberated speech, but the effect of phase distortion was still audible ((like a bell chime sounded at the background – non minimum-phase character of the acoustic impulse response – Chapter 2). There seemed to be no difference of phase audibility when compared to that of complete magnitude equalization.

**- For the 2nd case ($i = 1$):**

The effect of magnitude distortion was also perceived as significantly removed from the reverberated speech in this case, but the effect of phase distortion was less audible than in the first case.

**- For the 3thd case ($i = 2$):**

The effect of magnitude distortion was perceived as significantly reduced in this case as compared to that of the original signal (reverberated speech). When listening alternatively to both signals, it sounded like when switching between binaural (reverberated speech) and monaural (processeded signal) listening. The effect of phase distortion was not audible in this case. This case was subjectively preferred as compared to the previous cases.

The error criterion was also evaluated for both a second acoustic impulse response which is different to the initial one, and its corresponding equalized response using the same reduced complexity inverse filter used to equalize the initial one (Figure 3-5, (c) $i = 2$ ); assessing in this way potential robustness of the proposed method to possible changes in the listener position inside the room.

The results for magnitude equalization are tabulated below in the form of the objective error criterion evaluated for both original and equalized responses. The magnitude responses of the two cases (initial and second) for both original and equalized responses are shown in Figure 3-6.

**Table 3.2:** Objective error criterion (dB) evaluated for both original and equalized responses (Initial and second acoustic responses).

| **Acoustic Response** | **Original Response** (dB) | **Equalized Response** (dB) |
|---|---|---|
| **Initial** | 8.31 | 3.54 |
| **Second** | 7.46 | 5.26 |

These results indicate a significant reduction of the objective error criterion as it can be seen as spectral magnitude deviation from flatness for both cases (Figure 3-6). The subjective testing was also evaluated by listening tests through headphone (Logitech – Premium Stereo Headset) using Matlab`s Audio Tools and a reverberated anechoic speech (obtained by convolution of anechoic speech with the acoustic impulse response of Figure 3-1 (initial) and that corresponding to magnitude response of Figure 3-6-2 – (a) (second) respectively) as the original signals test. The subjective results indicated a significant perceptual preference for the equalized signal over the original for both cases (initial and second). These results can be considered as promising for the case of possible changes in the listener position inside the room.

**Summary**

FIR inverse filters have been designed for partial dereverberation in the context of an audioconferencing room. These are successively produced after application of an iterative simple complex smoothing operation to the corresponding measured acoustic impulse response. The most important results are:

- If the intention is to reduce reverberation, an FIR inverse filter such as that produced after three (3) iterations of the smoothing process (Figure 3-5, (c) $i = 2$) simplify the problem;
- if the intention is to remove reverberation, an FIR inverse filter such as that produced after one (1) iteration or two (2) iterations of the smoothing process (Figure 3-5, (a) $i = 0$, (b) $i = 1$) can resolve the problem at least for an objective single-point equalization.

## 3.5 CONCLUSIONS

In this chapter, equalization filters design based on iterative simple complex smoothing has been proposed for large rooms equalization. The algorithm can be useful in cases of long duration acoustic impulse responses since it results in impulse responses of reduced complexity which preserve the initial transient data that are significant to the listener. The algorithm also gives the possibility of stopping the smoothing process when the corresponding equalization filter produces some pre-determined equalized magnitude response – partial dereverberation.

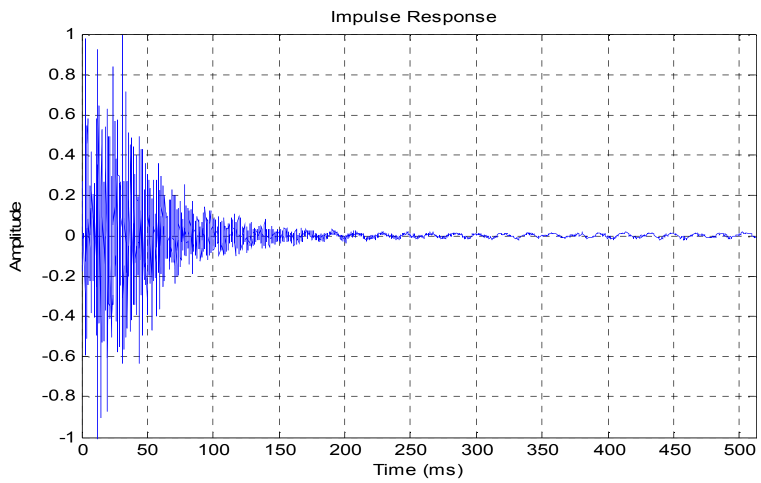



FIGURE 3.1**:** Acoustic impulse response measured in an audio-conferencing room

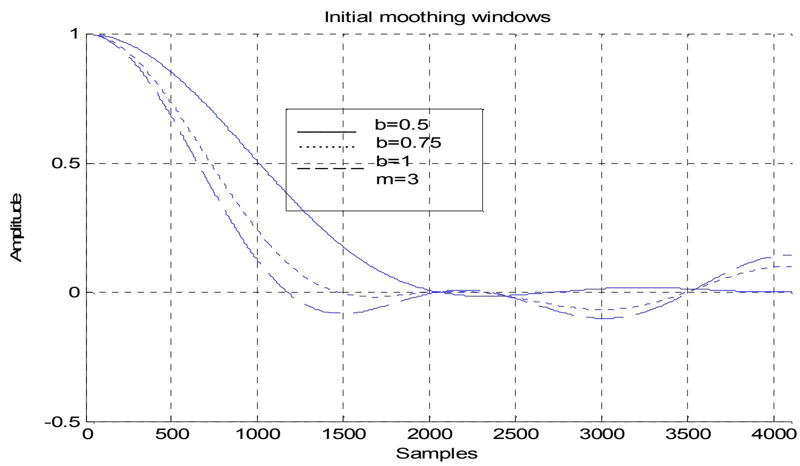



FIGURE **3**.2: Initial smoothing windows for $i = 0$

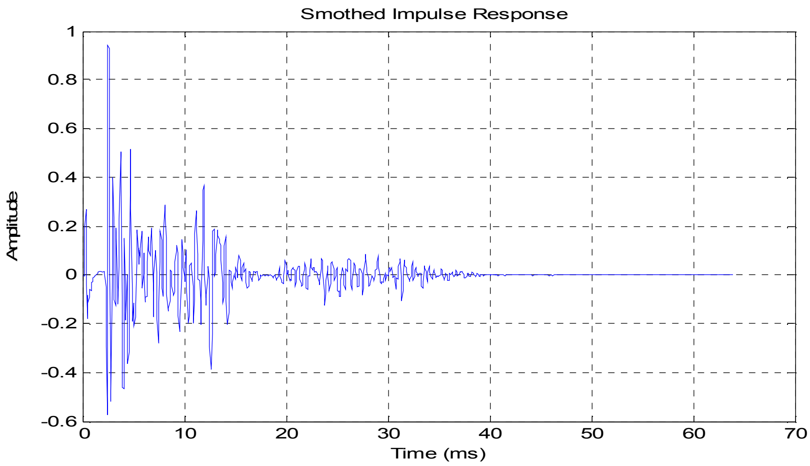



FIGURE **3**.3: Smoothed impulse response $h_{cs}^{3}(n),\ i=2$

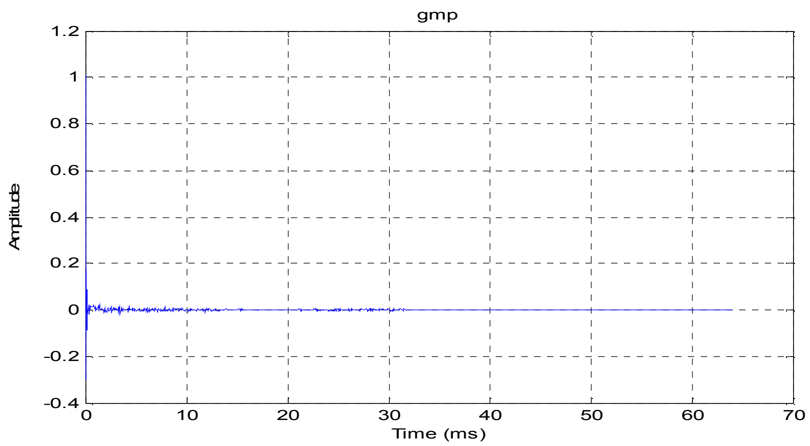
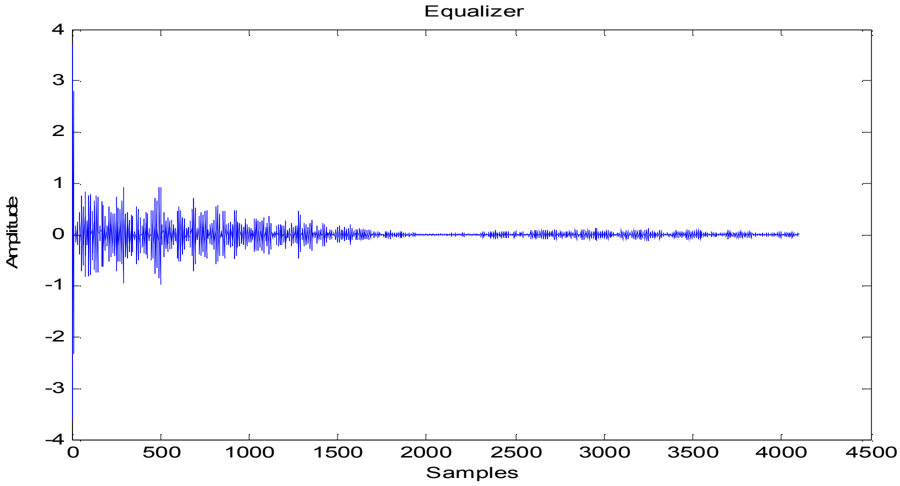
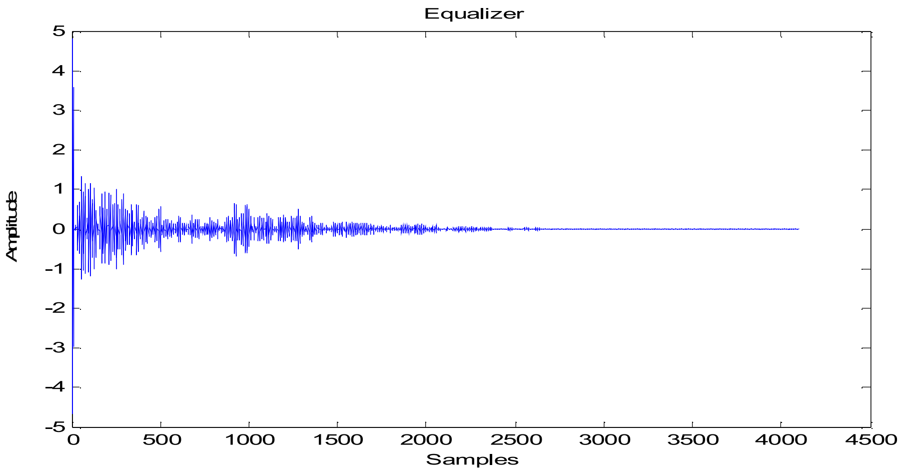


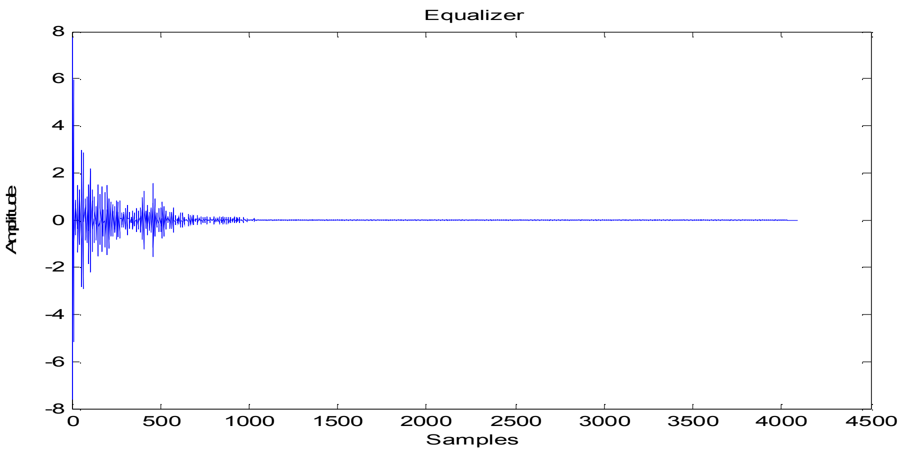

FIGURE **3**.4: Partial minimum-phase inverse impulse response $g_{cs,mp}^{1/8}(n)$ calculated using the iterative version of the homomorphic method with $L = 8$.




(a)

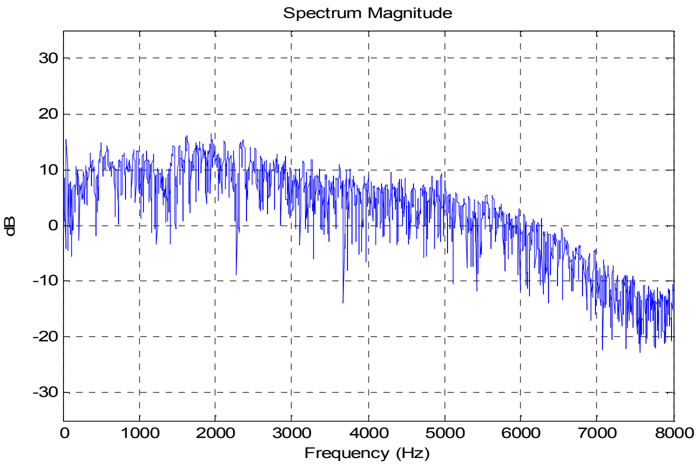



(b)

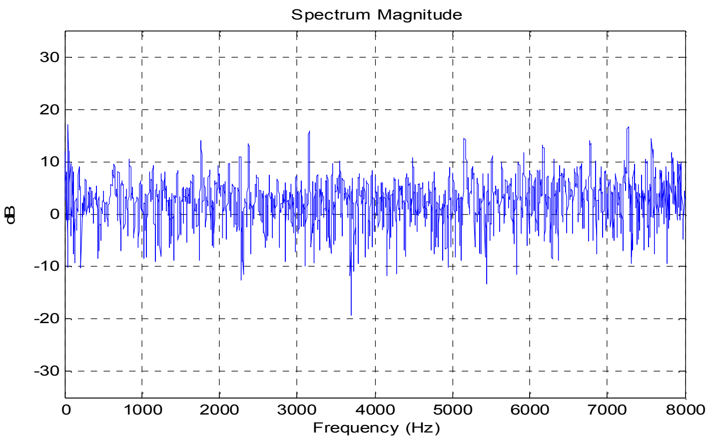



(c)

FIGURE 3.5: Impulse responses of the equalization filters obtained after 8 successive convolutions of partial minimum- phase inverse impulse responses - for : (a) $i = 0$, (b) $i = 1$ and (c) $i = 2$

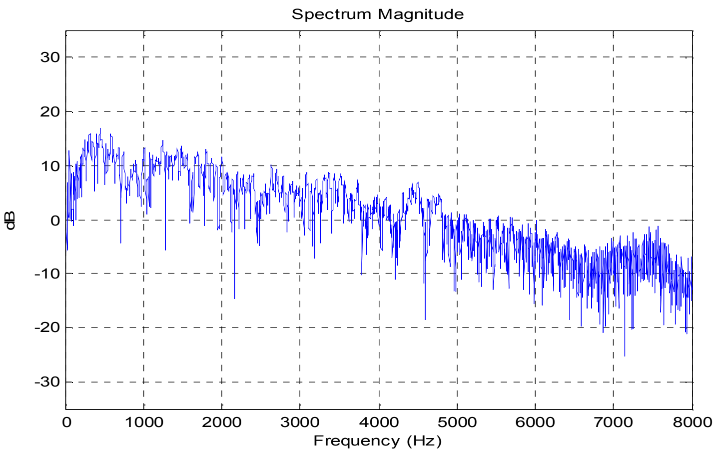



(a)

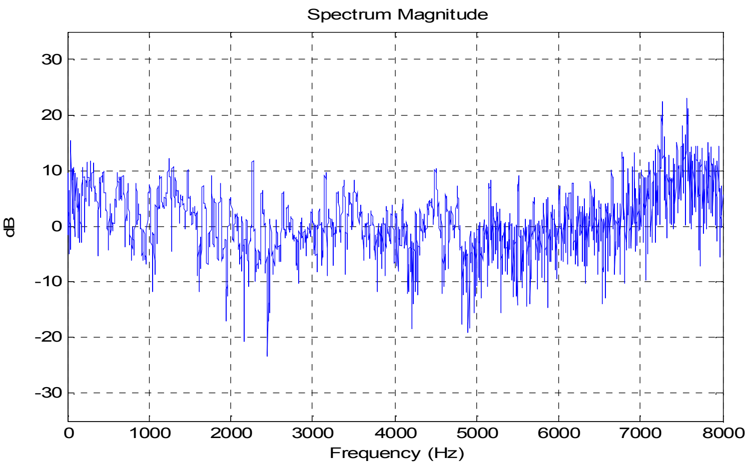



(b)

FIGURE **3**.6 – 1: Deviation of magnitude response from flatness
(a) Initial original response, b) Equalized response




(a)




(b)

FIGURE **3**.6 – 2: Deviation of magnitude response from flatness
a) Second original response, b) Equalized response

## 3.6 REFERENCES

[1] B. D. Radlovic and R. A. Kennedy, “Non-minimum phase equalization and its subjective importance in room acoustics”, IEEE Transactions on Speech and Audio Processing, vol. 8, no. 6, Nov 2000.

[2] J. N. Mourjopoulos, "Digital Equalization of Room Acoustics", J. Audio. Eng. Soc, Vol 42, N°11, 1994 November.

[3] P.Hatziantonio and J.Mourjoupoulos, " Generalized Fractional-Octave Smoothing of Audio and Acoustic Responses", J. Audio. Eng. Soc, vol 48, No 4, pp 259-280, 2000 April.

[4] P. M .Clarkson, J.Mourjopoulos and J.K.Hammond, "Spectral, Phase and Transient Equalization for Audio Systems", J. Audio. Eng. Soc, vol 33, pp 127-132, 1985 Mar.

[5] S. Salamouris, K. Politopoulos, V. Tsakiris and J. Mourjopoulos, “Digital System for Loudspeaker and Room Equalization ", presented at the 98th Convention of the Audio Engineering Society, J. Audio. Eng. Soc, (Abstracts), vol 43, pp 396, 1985 May.

[6] Neely S. T. and J. B. Allen, “Invertibility of a room impulse response”, J. Acoust. Soc. Am, vol. 66, no. 1, pp. 165-169, July 1979.

[7] B. D. Radlovic and R. A. Kennedy, “Iterative cepstrum-based approach for speech dereverberation ", ISSPA` 99, Brisbane, Australia, 22-25 August 1999.

[8] J.P.Jullien, A.Gilloire and A.Saliou, “Mesure des réponses impulsionnelles en acoustique”, Note technique NT/LAA/TSS/181, Novembre 1984

[9] **M. Ahfir**, I. Kale, A. Krukowski and D. Berkani, ‘’ Room Equalization Based on Iterative Simple Complexe Smoothing of Acoustic Impulse Responses ’’, the 12th International Conference ICECS’05, TUNISIA, Vol 1, pp 256-259, Dec 2005.

Chapter 4

# RESULTS VALIDATION

This Chapter presents some preliminary simulation results obtained when the smoothing process has been applied for a dereverberation system placed before a simple speech recognition system of isolated words with mono-speaker, and listening tests validation for a large audio-conferencing room after implementation of an equalization filter example on Texas Instruments development board. For the speech recognition system we choosed the equalization filters corresponding to the first and second iteration. Because reverberation is significantly removed in these two cases (simulations results in Chapter 3) which is conform to such application, where speech must be as clean as possible. For the listening tests evaluation we choosed the equalization filter corresponding to the third iteration. Because reverberation in this case is significantly reduced which make speech sound naturally in the room.

## 4.1 APPLICATION OF SMOOTHING PROCESS TO A SPEECH RECOGNITION SYSTEM OF ISOLATED WORDS

Current speech recognition systems almost universally assume noiseless speech input and have been trained with and expect the incoming signal to be free of any ambient and interfering noise. Typical solutions to achieve this have been requiring the speaker to be in an acoustically treated room, restricting ambient noise, or using special microphones. There has been great interest in using a desktop microphone as the input device to the recognition system instead of the unnatural headset microphone. Unfortunately, by making the input device more natural for the user, the word recognition rate of the system can fall dramatically. Most of the degradation is thought to be caused by the acoustical characteristics of the recording environment, particularly the reverberation [1]. Perfect dereverberation of the signal is an ideal solution, but as it is known, this requires very long inverse filters length for implementation. The goal of this application is to evaluate how a partial dereverberation can improve the performance of a speech recognition system in a reverberant environment assuming the training to be in an anechoic chamber. In the next section, we present some preliminary results of the previous technique (iterative simple complex smoothing of acoustic impulse responses) application in the case of a simple speech recognition system of isolated French words, and with mono-speaker.

## 4.2 RESULTS

The smoothing process was also applied to a simple speech recognition system of two isolated different French words, and with mono-speaker (Figure 4-1), [2]. This simple case was considered to assess the effect of smoothing process when applying a speech dereverberation system (partial single-point equalization) before the recognition system. The most used decision criterion in isolated words recognition known as *Euclidian Distance* was adopted for this assessment and is given as follows:

$$d(a,b) = \left[ \sum_{k=1}^{p} |a_k - b_k|^2 \right]^{1/2} , \qquad (4.1)$$

where $a$ is a word vector of references, recorded in this application in an anechoic chamber to make speech perfectly noiseless (training), and $b$ is the incoming word vector to be recognized. $p$ (equal between 8 and 14) being the model order of LPC (linear prediction coding) parameters of $a$ and $b$ respectively. The objective is to minimize the distance $d(a,b)$ by comparison of the word $b$ with all the references. This means that a reference word $a$ will be identified which is the nearest copy of the word $b$ .

In this application the previous acoustic impulse response of an audio-conferencing room (where the reverberation phenomenon is more significant - figure 3-1) was used to generate reverberated words by convolution with anechoic words. The equalized words were generated by application of equalization filters to reverberated words according to the smoothing process Figure 3-5-(a) $i=0$ and Figure 3-5-(b) $i=1$, respectively, and where reverberation is considered to be significantly removed after these two iterations. The *Euclidian Distance* $d(a,b)$ was evaluated for anechoic, reverberated and equalized words respectively. The results in the case of two ( $N=2$ ) French words as references are tabulated below.

Firstly, these results indicate that the two different words are identified even in reverberant environment as in the ideal case in anechoic environment, with some degradation of the Euclidian Distance. But this is a limited word recognition system ($N = 2$). This means that such degradation can cause a dramatic fall to the word recognition rate of the system in case of a high number $N$ of words. When applying a dereverberation system (equalization filters) before the recognition system which is based on smoothing process of the corresponding acoustic impulse response for one (Figure 3-5- (a) $i = 0$) and two iterations (Figure 3-5- (b) $i = 1$) respectively, we can see an improvement of the performance function of the number of iteration and then FIR inverse filters complexity. These results can be considered as preliminary for a future application, where the number of isolated words will be very high.

**Table 4.1:** *Euclidian Distance* evaluated for anechoic, reverberated and equalized words

| Words | Word 1 | Word 2 |
|---|---|---|
| Anechoic words | 0 | 0 |
| Reverberated words | 0.6527 | 0.4261 |
| Equalized words for two iteration of smoothing process, $i = 1$ | 0.5119 | 0.2095 |
| Equalized words for one iteration of smoothing process, $i = 0$ | 0.4304 | 0.2073 |

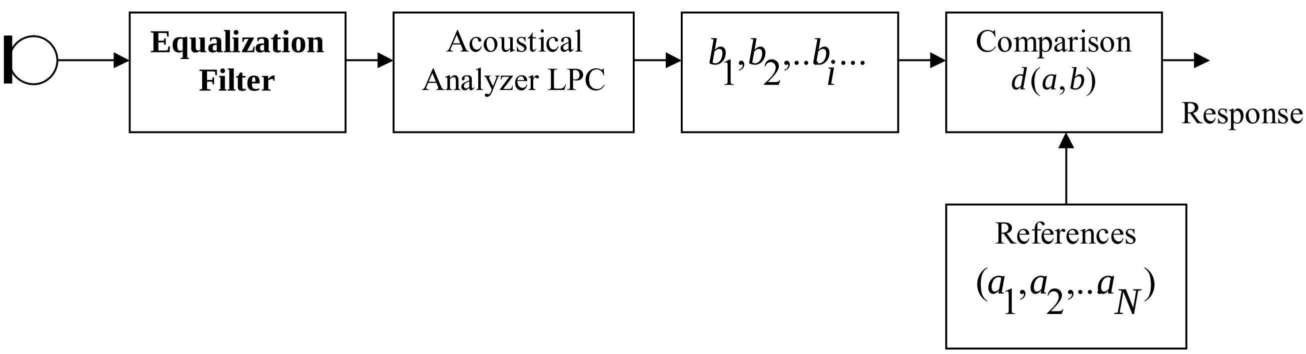



FIGURE 4-1: Speech recognition system of isolated words

## 4.3 IMPLEMENTATION OF AN EQUALIZATION FILTER ON TMS320VC5510 DSK DSP PLATFORM

**EXPERIMENT**

An equalization filter representing the example of Chapter 3 (Figure 3-5, (c) $i = 2$) with reduced complexity (significantly about 1000 samples rather than more than 10000 required in a complete equalization – perfect dereverberation) has been implemented on TMS320VC5510 DSK kit (Figure 4-2). This impulse response corresponds to 3.54 dB and 5.26 dB as objective error criterion for the initial and second equalized impulse responses respectively (Table 4-2). We used a version 2 of Code Composer Studio to load on the DSP platform the C source coefficients (coded 16 bits integer – int16) of the filter generated by MATLAB and execute the C source program of an FIR filter implementation (see APPENDIX B – from help instructions of version 2 of Code Composer Studio). The test signals (wave files coded 16 bits at 16 kHz) transmitted from PC used in this application for listening evaluations were the same used in chapter 4, but passed trough AIC23 Codec of the DSP kit (for processing from the line input (LINE IN) to the line output (LINE OUT)). If DIP switch 3 is high no filtering of audio Data, but if DIP switch 3 is down audio Data are filtered by the FIR coefficients before playing the results from LINE OUT of AIC23 Codec.

The first signal is a reverberated speech, obtained after convolution of an anechoic speech with an initial acoustic impulse response of an audio-conferencing room (Figure 3-1 and corresponding to magnitude response of Figure 3-6-1- (a)). The second signal is also a reverberated speech, but obtained after convolution of the same anechoic speech with a second acoustic impulse response which is different than the initial one (corresponding to magnitude response of Figure 3-6-2 – (a)), representing in this way possible changes in the listener position inside the room. The third and fourth signal are respectively the equalized (filtered) first and second signal (first and second signal passed trough DSP module for filtering by FIR coefficients - when DIP switch 3 is down - before playing the results from the output (LINE OUT) of AIC23 Codec).

These four (04) test signals recorded by MATLAB's Audio Tools from LINE OUT of AIC23 Codec were presented binaurally in randomly order to young listeners (about

25 graduation students in a quiet listening class room) with normal hearing through a headphone (Logitech – Premium Stereo Headset) at a comfortable level. The listening tests evaluation indicates in Figure 4-3 as it was expected in Chapter 3, a significant perceptual preference for the equalized signals over the original signals (reverberated speech) for both cases, the first and second signal (that is, respectively the third and the fourth signal). These results can be confirmed in an **anechoic chamber** by transmitting the recorded wave files trough an adapted loudspeaker to some listeners placed inside.

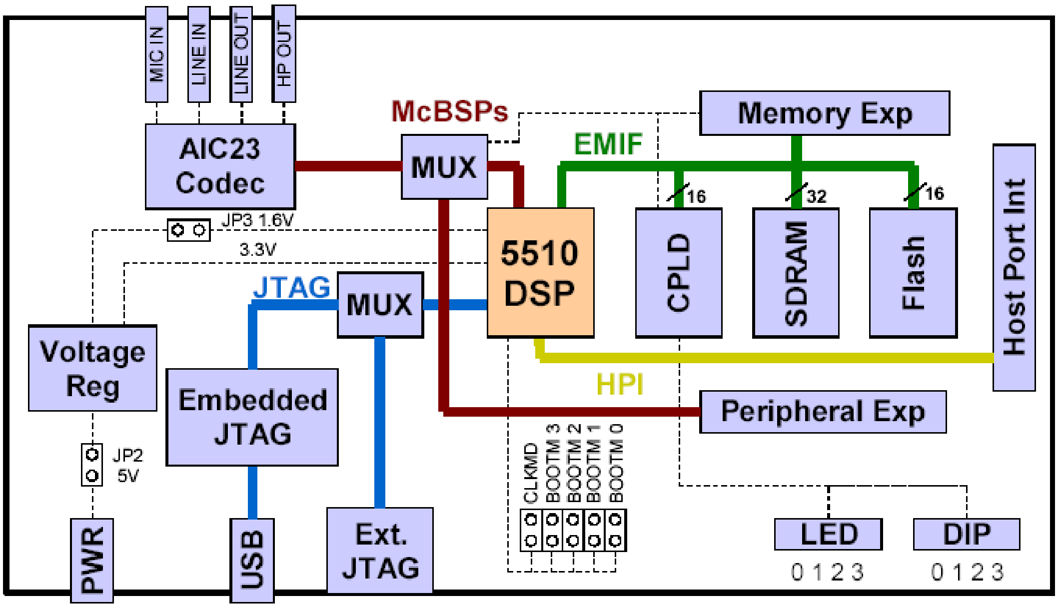



FIGURE 4-2: Block diagram VC5510 DSK

**Key Features:**

• A Texas Instruments 5510 DSP operating at 200MHz

• An AIC23 stereo codec

• 8 Mbytes of synchronous DRAM

• 512 Kbytes of non-volatile Flash memory

• 4 user accessible LEDs and DIP switches

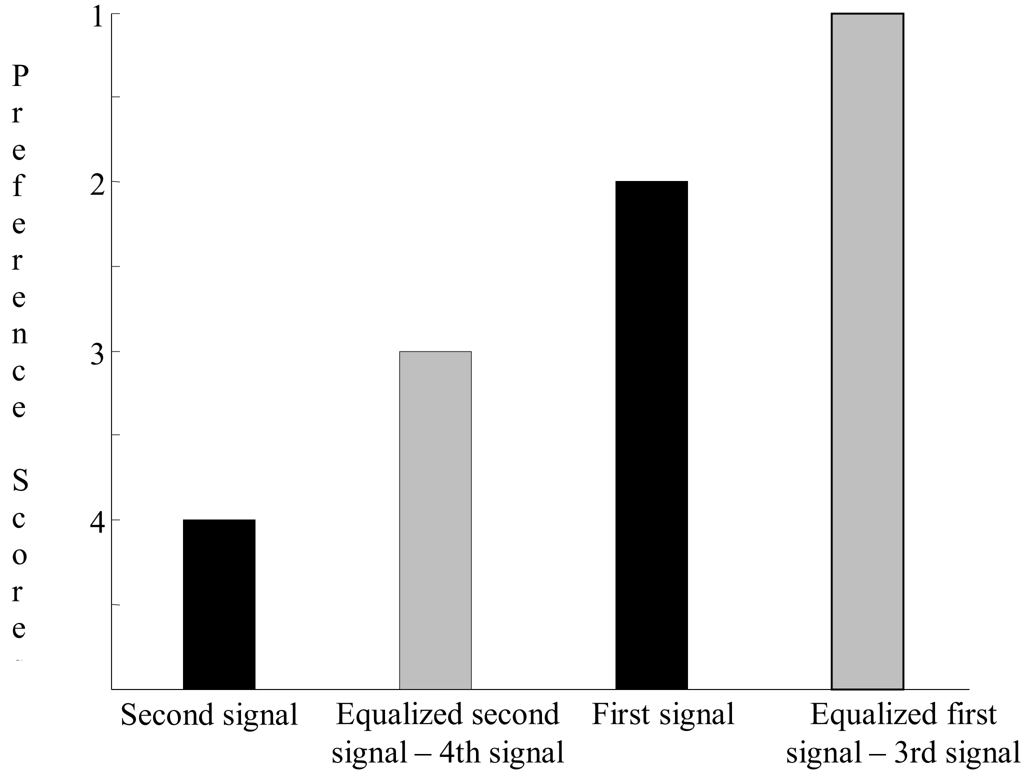
• Software board configuration through registers implemented in CPLD

• Jumper selectable boot options

• Standard expansion connectors for daughter card use

• JTAG emulation through on-board JTAG emulator with USB host interface or external emulator

• Single voltage power supply (+5V)




Preference Score 1: the most preferred signal

FIGURE 4-3 Listening tests evaluation

**4.4 CONCLUSIONS**

Reduced complexity FIR Inverse filters have been applied before a simple speech recognition system. Simulation results are preliminary but promising for a high number of words identification. Evaluation of listening tests carried out on Texas Instrument DSP Board for a reduced complexity FIR inverse filter corresponding to reduced reverberation confirmed that reported in Chapter 3 by simulation with MATLAB's Audio Tools.

**4.5 REFERENCES**


[1] S.R.Palm, "Enhancement of Reverberated Speech", Master of Science at Carnegie Mellon University, Pittsburgh, Pennsylvania, 15213-3890.

[2] CALLIOPE, "La parole et son traitement automatique", MASSON, Paris, Milan, Barcelone, Mexico 1989.

## CONCLUSIONS AND FUTURE WORK

Standard homomorphic method to design stable minimum-phase inverse filters for non-minimum-phase acoustic impulse responses equalization has been presented in detail.

A modified version of the standard homomorphic method for small rooms equalization has been proposed. This version is useful in case of partial magnitude equalization, where the dominant zeros density of the system is not very high. Although it is used in this work as an additional optimizing tool for the psycho acoustic quality measurement of speech, this alternative approach is advantageous in case of the direct inverse filtering (minimum-phase system) when perfect equalization of a small reverberant room is not desired. Simulation Results using objective and subjective measurements for an impulse response measured in the car interior show that by using the modified version we can control the sound quality more precisely than when using the standard method.

In case of large rooms, an equalization filters design method has been proposed. This method is based on a modification of the corresponding long duration measured acoustic impulse responses by using an iterative simple smoothing process. Corresponding impulse responses of reduced complexity are then produced, but preserve the initial transient data that are significant to the listener (direct signal and first reflections). The algorithm also gives the possibility of stopping the smoothing process when a corresponding designed equalization filter (inverse filter) produces some pre-determined equalized magnitude response. Equalization results have been validated in the context of an audio-conferencing room for a reduced complexity equalization filter (inverse filter) using listening test with a headphone, MATLAB's Audio Tools and implementation on Texas Instruments DSP board.

Finally, partial dereverberation can be applied to improve the performance of a speech recognition system, in case of isolated words identification with mono-speaker.

In the future work, partial dereverberation will be applied to phonemes recognition systems with multi-speakers.

# APPENDIX A

Proof for the relation (3-13), Chapter 3:

The real part of the complex cepstrum of *h(n)* in the equation (3-6) is defined as the inverse DFT of the function

$$\breve{H}(k) = \log\left|H(k)\right| . \qquad \text{A.1}$$

Applying the direct DFT on the real cepstrum of the minimum-phase $h_{mp}(n)$ in the relation (3-7) for $L=1$, leads to

$$\breve{H}_{mp}(k) = \log\left|H_{mp}(k)\right| . \qquad \text{A.2}$$

For $L \neq 1$, the relation A.2 becomes

$$\breve{H}_{mp}(k) = \frac{1}{L}\log\left|H_{mp}(k)\right| . \qquad \text{A.3}$$

Using the relation (3-3), the minimum-phase part $H_{mp}(k)$ of the relation (3-9) can be expressed as follows:

$$H_{mp}(k) = \exp(\frac{1}{L}\log\left|H_{mp}(k)\right|) = \exp(\frac{1}{L}\log\left|H(k)\right|) \qquad \text{A.4}$$

Therefore, the inverse of $H_{mp}(k)$ , $G_{mp}(k)$ is given by

$$G_{mp}(k) = \exp(-\frac{1}{L}\log\left|H(k)\right|) , \qquad \text{A.5}$$

or

$$\log\left|G_{mp}(k)\right| = -\frac{1}{L}\log\left|H(k)\right| . \qquad \text{A.6}$$

# CONTRIBUTIONS

Hindawi Publishing Corporation
EURASIP Journal on Applied Signal Processing
Volume 2006, Article ID 67467, Pages 1–8
DOI 10.1155/ASP/2006/67467

# Partial Equalization of Non-Minimum-Phase Impulse Responses

**Ahfir Maamar,[1] Izzet Kale,[2,3] Artur Krukowski,[2,4] and Berkani Daoud[5]**

[1] *Department of Informatics, University of Laghouat, BP 37G, Laghouat 03000, Algeria*
[2] *Department of Electronic Systems, University of Westminster, 115 New Cavendish Street, London W1W 6UW, UK*
[3] *Department of Electronic Systems, Eastern Mediterranean University, Gazimagusa, Mersin 10, Cyprus*
[4] *National Center of Scientific Research "Demokritos," Agia Paraskevi, Athens 153 10, Greece*
[5] *Department of Electronics, Ecole Nationale Polytechnique, BP 182, Algiers 16000, Algeria*

Received 1 March 2005; Revised 5 December 2005; Accepted 26 February 2006

Recommended for Publication by Piet Sommen

We propose a modified version of the standard homomorphic method to design a minimum-phase inverse filter for non-minimum-phase impulse responses equalization. In the proposed approach some of the dominant poles of the filter transfer function are replaced by new ones before carrying out the inverse DFT. This method is useful when partial magnitude equalization is intended. Results for an impulse response measured in the car interior show that by using the modified version we can control the sound quality more precisely than when using the standard method.



## 1. INTRODUCTION

In sound-reproduction systems an equalization filter is often used to modify the frequency spectrum of the original source before feeding it to the loudspeaker. The purpose is to make the impulse response of the equalized sound-reproduction chain as close as possible to the desired one [1]. In principle the direct inversion of mixed-phase (or non-minimum-phase) measured impulse responses of the systems is not possible since it leads to unstable equalization filter realizations. Since any mixed-phase impulse response can be represented mathematically by the convolution of a minimum-phase sequence and a maximum-phase (or all-pass) sequence [2], it is possible to derive and implement an approximate and stable inverse filter for such systems [3]. This is because a causal and stable sequence can invert the minimum-phase component of any mixed-phase sequence and an infinite acausal (anticipatory) and stable sequence can similarly invert the maximum-phase component of such sequences [3]. For the reason of the implementation complexity of such combined equalization filters as it will be discussed in Section 2, the work presented in this paper focuses on the equalization of the minimum-phase component of the system and its partial equalization importance. One method to design such a minimum-phase equalization filter is the homomorphic one based on the measured impulse response of the system. This method known as standard used for the case of single-point equalization is described in Section 2. In Section 3, a modified version of the standard homomorphic method is proposed. It takes into account that the listener is able to detect gradual response variations of less than 0.5 dB [4, 5] and hence is able to control the sound quality more accurately. Section 4 shows the magnitude equalization performance results for an impulse response measured in a car interior using both objective and subjective measurements.

## 2. STANDARD HOMOMORPHIC METHOD

A non-minimum-phase discrete impulse response, $h(n)$, of a system can be described as [2]

$$h(n) = h_{\mathrm{mp}}(n) \otimes h_{\mathrm{ap}}(n), \tag{1}$$

where $\otimes$ denotes the discrete convolution. This can be shown in the frequency domain as

$$H(k) = H_{\mathrm{mp}}(k) H_{\mathrm{ap}}(k), \tag{2}$$

where $h_{\mathrm{mp}}(n)$ is a minimum-phase sequence, such that its DFT, $H_{\mathrm{mp}}(k)$, satisfies the relation

$$|H_{\mathrm{mp}}(k)| = |H(k)|, \tag{3}$$

where $H(k)$ is the DFT of $h(n)$ given by

$$H(k) = \sum_{n=0}^{N-1} h(n)e^{-j(2\pi kn/N)}, \quad (4)$$

where $N$ is the length of $h(n)$ and $h_{ap}(n)$ is an all-pass sequence of $|H_{ap}(k)| = 1$, for $k = 0, 1, \dots, N-1$.

The convolution operation of $h_{mp}(n)$ and $h_{ap}(n)$ can be expressed as the algebraic addition of their corresponding complex cepstra $\hat{h}_{mp}(n)$ and $\hat{h}_{ap}(n)$ by the homomorphic transformation [6]. This leads to a decomposition of a non-minimum-phase impulse response into its minimum-phase and all-pass components. The standard homomorphic method algorithm is outlined as follows [4, 7, 8].

(1) Compute the DFT of $h(n)$.

(2) Compute

$$\hat{H}(k) = \log |H(k)|. \quad (5)$$

(3) Compute the real part of the complex cepstrum of $h(n)$,

$$\hat{h}(n) = \frac{1}{N}\sum_{k=0}^{N-1} \log |H(k)| e^{j(2\pi kn/N)}, \quad (6)$$

for $n = 0, 1, \dots, N-1$.

(4) Compute the corresponding real cepstrum of the minimum-phase $h_{mp}(n)$,

$$\hat{h}_{mp}(n) = \begin{cases} \dfrac{\hat{h}(n)}{L}, & n = 0, \dfrac{N}{2}, \\ \dfrac{2\hat{h}(n)}{L}, & 1 \le n \le \dfrac{N}{2}, \\ 0, & \dfrac{N}{2} < n \le N-1, \end{cases} \quad (7)$$

where $L$ is a positive real parameter [8].

(5) Compute the DFT of $\hat{h}_{mp}(n)$,

$$\hat{H}_{mp}(k) = \sum_{n=0}^{N-1} \hat{h}_{mp}(n)e^{-j(2\pi kn/N)}. \quad (8)$$

(6) Compute *the minimum-phase part* $H_{mp}(k)$,

$$H_{mp}(k) = \exp\left(\hat{H}_{mp}(k)\right). \quad (9)$$

(7) Compute the equalized response, $H_{eq}(k)$,

$$H_{eq}(k) = H(k)G_{mp}(k), \quad (10)$$

where $G_{mp}(k)$ represents the inverse of $H_{mp}(k)$,

$$G_{mp}(k) = \frac{1}{H_{mp}(k)}. \quad (11)$$

In the time domain, this is equivalent to a deconvolution,

$$h_{eq}(n) = h(n) \otimes g_{mp}(n), \quad (12)$$

with $g_{mp}(n)$, being the inverse DFT of $G_{mp}(k)$.

In the case of $L = 1$, the algorithm corresponds to a magnitude equalization. If a sufficiently large number $N$ is used for DFT computation, the effect of magnitude distortion caused by the system can be perfectly removed in practice by convolving $h(n)$ with the inverse minimum-phase impulse response $g_{mp}(n)$ [3, 7]. The effect of phase distortion can also be solved by convolving the all-pass sequence, $h_{ap}(n)$, (obtained after deconvolution of $h(n)$ with $g_{mp}(n)$) with its time reversed version, $h_{ap}(-n)$, [4, 9]. As a result, implementation of such combined equalization (complete equalization) requires very long FIR filters. But this is not always required in practice. For this reason, the equalization of the all-pass component (phase equalization) will not be considered in this work.

In the case of $L > 1$, the algorithm corresponds to a partial magnitude equalization. This requires a shorter FIR filter to keep the phase distortion below the threshold of audibility [4].

Sometimes, the frequency response of the system, $H(k)$, and hence its minimum-phase part, $H_{mp}(k)$, can be represented by a low number of isolated dominant zeros. In such a case, increasing the parameter $L$ by a significant value during the control process may shorten the length of the equalization filter too, resulting in an unsatisfactory equalization performance. This is because an increase in $L$ results in a decrease of all the radii of the complex poles of $G_{mp}(k)$ together according to the relation derived from (3), (7), and (9) (see the appendix),

$$\log |G_{mp}(k)| = -\frac{1}{L}\log |H(k)|. \quad (13)$$

This means that the complex poles of $G_{mp}(k)$ appear to be pushed together towards the origin of the unit circle.

In the next section, we propose an alternative approach in which, instead of pushing all the poles of $G_{mp}(k)$, we push the most dominant of them selectively and slightly towards the origin of the unit circle by decreasing the corresponding high values of the $Q$ factors (values of the steady-state resonances). This allows controlling the magnitude equalization performance more precisely—especially applicable in practice. The reason is that the listener is able to detect gradual response variations of less than 0.5 dB [4, 5]. Furthermore, the proposed technique is advantageous when the parameter $L$ cannot be calculated theoretically, for example, for the case of the direct inverse filtering (no cepstral analysis, i.e., no steps 2 to 6) of a small reverberant room where the dominant poles can be identified even if they are closely spaced [10, 11].

## 3. MODIFIED VERSION

A replacing method of some dominant poles of the inverse minimum-phase function $G_{mp}(k)$ is described in this section. They are identified using the standard method ($L = 1$)

and then replaced before doing the inverse DFT in order to calculate the corresponding discrete time sequence $\tilde{g}_{\text{mp}}(n)$ representing the impulse response of the new equalization filter.

The $z$ transform function of a complex pole pair is expressed as [10, 12]

$$H_p(z) = \frac{1}{(1 - |a|e^{j\theta}z^{-1})(1 - |a|e^{-j\theta}z^{-1})},$$

or (14)

$$H_p(z) = \frac{1}{1 - 2|a|\cos\theta z^{-1} + |a|^2 z^{-2}},$$

where $|a|$ is the pole radius in the $z$ plane and $\theta = 2\pi(f_p/f_e)$ is its phase angle with $f_e$ being the sampling frequency and $f_p$ the frequency of the complex pole.

Taking the inverse $z$ transform of $H_p(z)$ the corresponding impulse response is [10, 12]

$$h_p(n) = \frac{|a|^n \sin(n\theta + \theta)}{\sin(\theta)} u(n), \tag{15}$$

where $u(n)$ is a unit step function.

The transfer function of the selective filter for a complex pole pair is

$$H_s(z) = \frac{\tilde{H}_p(z)}{H_p(z)}, \tag{16}$$

where the transfer function $\tilde{H}_p(z)$ contains a new complex pole pair at the same frequency of the old pair but at a desired smaller radius, $|\tilde{a}|$. This technique allows us to decrease selectively the $Q$ factors values of a low order of isolated pole pairs in the frequency response $G_{\text{mp}}(k)$. The new inverse minimum-phase function becomes

$$\tilde{G}_{\text{mp}}(z) = G_{\text{mp}}(z) H_s^{(1)}(z) \cdots H_s^{(P)}(z), \tag{17}$$

and its discrete version

$$\tilde{G}_{\text{mp}}(k) = G_{\text{mp}}(k) H_s^{(1)}(k) \cdots H_s^{(P)}(k), \tag{18}$$

where $P$ is the number of identified and replaced dominant pole pairs from $G_{\text{mp}}(k)$, and $H_s^{(P)}(k)$, $p = 1, \ldots, P$, are the sampled frequency responses of selective filters equal to the number of replaced pole pairs, $P$.

This function is then inverted using the inverse DFT in order to obtain its discrete time domain equivalent $\tilde{g}_{\text{mp}}(n)$, shorter than $g_{\text{mp}}(n)$ calculated by standard method for $L = 1$ and longer than $g_{\text{mp}}(n)$ obtained for $L > 1$.

One method to identify frequencies of the isolated poles is to iteratively search for the increased magnitude response level of $G_{\text{mp}}(k)$ caused by poles (peaks) residing within the frequency range of interest (in our case below 4 kHz). In each iteration a maximum magnitude level $G_{\text{mp}}(f_p)$ corresponding to the highest pole frequency $f_p$ is found. This technique was found robust even in the case of very closely spaced poles [10, 11]. After determining the frequency $f_p$ of the highest pole, the corresponding pole radius must be determined based on the $Q$ factor value according to the following relation, since our work here is restricted to a low order of isolated poles [10, 13–15],

$$Q = G_{\text{mp}}(f_p) = \frac{1}{1 - |a|}. \tag{19}$$

The replacing method means that the dominant poles of $G_{\text{mp}}(k)$ are identified one by one and then replaced iteratively by new ones, where each corresponds to a desired $\tilde{Q}$ factor, $\tilde{Q} = 1/(1 - |\tilde{a}|)$, starting from the most dominant one.

The implementation algorithm of the proposed modified method (useful for partial magnitude equalization) is as follows.

(1) Compute the steps 1 to 6 as in the standard method for $L = 1$.
(2) Compute the inverse minimum-phase

$$G_{\text{mp}}(k) = \frac{1}{H_{\text{mp}}(k)}, \tag{20}$$

using $G_{\text{mp}}(k) = \tilde{G}_{\text{mp}}(k)$.
(3) Set $p$ to 1.
(4) Estimate the most dominant pole from $\tilde{G}_{\text{mp}}(k)$ as described above, (determine $f_p$ and $|a|$).
(5) Design its selective filter using (16).
(6) Replace the estimated pole from $\tilde{G}_{\text{mp}}(k)$ using (18).
(7) Increment $p = p + 1$ and repeat the steps 4, 5, and 6 until $p = P$.
(8) Compute $\tilde{g}_{\text{mp}}(n)$ as inverse DFT of $\tilde{G}_{\text{mp}}(k)$.
(9) Compute the equalized response $H_{\text{eq}}(k)$,

$$H_{\text{eq}}(k) = H(k)\tilde{G}_{\text{mp}}(k). \tag{21}$$

In the time domain, this is equivalent to a deconvolution

$$h_{\text{eq}}(n) = h(n) \otimes \tilde{g}_{\text{mp}}(n). \tag{22}$$

In the next section we present the performance evaluation of the magnitude equalization performed by the proposed version as compared to that from standard method, using both objective measures based on an error criterion and subjective tests of speech quality.

## 4. RESULTS

In order to assess the performance of our algorithms, we used a frequency domain error criterion, which estimates the standard deviation of the magnitude response from a constant level [4]. The error criterion $\Delta$(dB) is given as follows:

$$\Delta = \left[\frac{1}{N}\sum_{k=0}^{N-1}\left(10\log_{10}\left|H_{\text{eq}}(k)\right| - H_m\right)^2\right]^{1/2}, \tag{23}$$

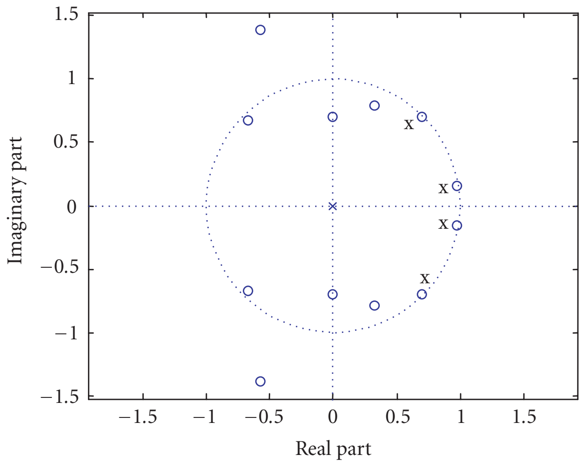



Figure 1: Complex $z$ plane of six known zeros (poles after inversion of the minimum-phase part): (o) pole pairs and (x) replacing pole pairs.

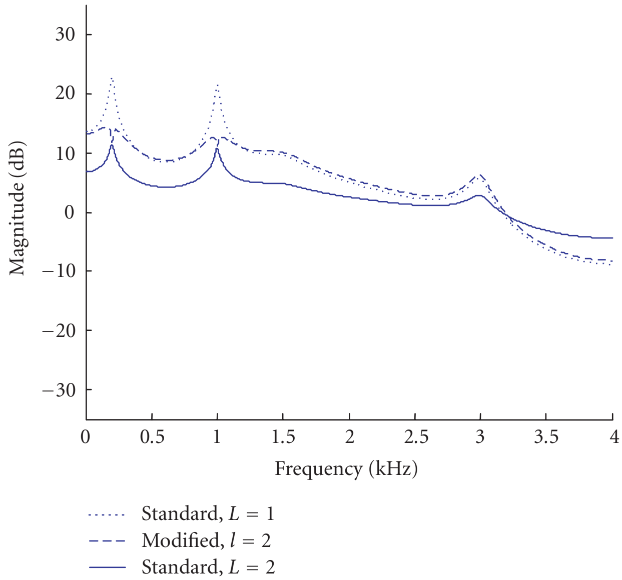



Figure 2: Inverse frequency response $G_{mp}(k)$ calculated using different methods.

where

$$H_m = \frac{1}{N}\sum_{k=0}^{N-1} 10\log_{10}|H_{eq}(k)|. \quad (24)$$

Two examples of non-minimum-phase impulse responses were used to compare both algorithms. The first one was synthetic, used just to enlighten the replacing approach of a low order of isolated poles, and the second one used real measurements taken in a car interior. We also introduced in the proposed version a real parameter $l$ ($l > 1$), in order to selectively decrease the highest $Q$ factors of dominant poles. This means that the new replacement poles correspond to desired $\widetilde{Q}$ factors, $\widetilde{Q} = Q/l$.

### 4.1. Synthetic impulse response

We first considered a simple synthetic impulse response. This was obtained by successive convolution of six known zeros sequences somewhat isolated in the complex $z$ plane (Figure 1) and with at least one placed outside the unit circle, which make the impulse response a non-minimum-phase one. These are defined as ($f_e = 8$ kHz):

$$\begin{aligned} |a| &= 0.99 \quad \text{at } f_p = 200\,\text{Hz},\\ |a| &= 0.99 \quad \text{at } f_p = 1000\,\text{Hz},\\ |a| &= 0.85 \quad \text{at } f_p = 1500\,\text{Hz},\\ |a| &= 0.70 \quad \text{at } f_p = 2000\,\text{Hz},\\ |a| &= 1.5 \quad \text{at } f_p = 2500\,\text{Hz},\\ |a| &= 0.95 \quad \text{at } f_p = 3000\,\text{Hz}. \end{aligned} \quad (25)$$

Figure 2 shows the inverse frequency response $G_{mp}(k)$ from which the two ($P = 2$) most dominant poles are estimated and corresponding to

$$\begin{aligned} |a_1| &= 0.9947 \quad \text{at } f_1 = 200\,\text{Hz},\\ |a_2| &= 0.9915 \quad \text{at } f_2 = 1000\,\text{Hz}. \end{aligned} \quad (26)$$

These two ($P = 2$) poles are selectively replaced by two new poles of smaller radius corresponding to $Q_1/l$ and $Q_2/l$ factors, respectively, with a significant value of $l$, ($l = 2$) (Figure 1). In Figure 2, even with some error in the estimation of poles, we still can observe a decrease in the $Q$ factors depending on the position of the new poles. This corresponds to a reduction of $\tilde{g}_{mp}(n)$ length when compared to $g_{mp}(n)$. When using the standard method and considering the same significant value of $L$ ($L = l = 2$), we can see in Figure 2 that all the poles have been pushed together towards the origin of the unit circle too, resulting also in the reduction of the $g_{mp}(n)$ length, but this is considerably shorter than that of $\tilde{g}_{mp}(n)$.

The evaluation of the objective error criterion for this example is not considered because of its little practical interest.

### 4.2. Practical impulse response

A real impulse response was measured for the car interior at a sampling frequency of 8 kHz. A record of 1024 samples was zero padded up to $N = 2048$. This impulse response is shown in Figure 3. In Figure 4 an unstable direct inverse impulse response is shown, demonstrating its non-minimum-phase character.

Figure 5 shows the inverse minimum-phase frequency response $G_{mp}(k)$. It was calculated using the standard method ($L = 1$). The most dominant pole can be clearly seen there. The search was limited to a single pole, that is, $P = 1$, such that $|a_1| = 0.9993$ at $f_1 = 70.38$ Hz that caused the inverse

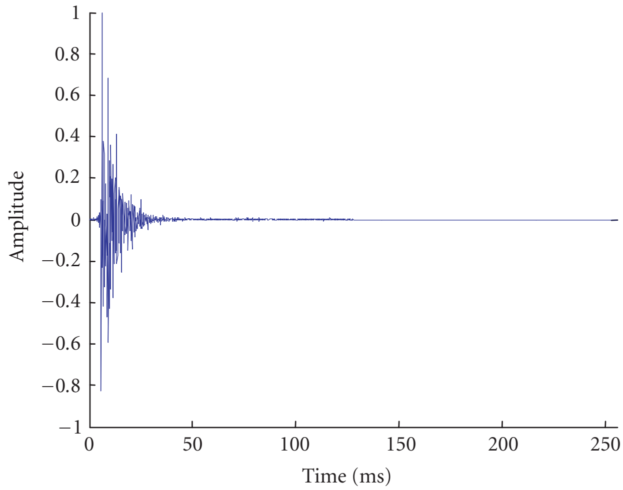



Figure 3: Impulse response measured in the car interior.

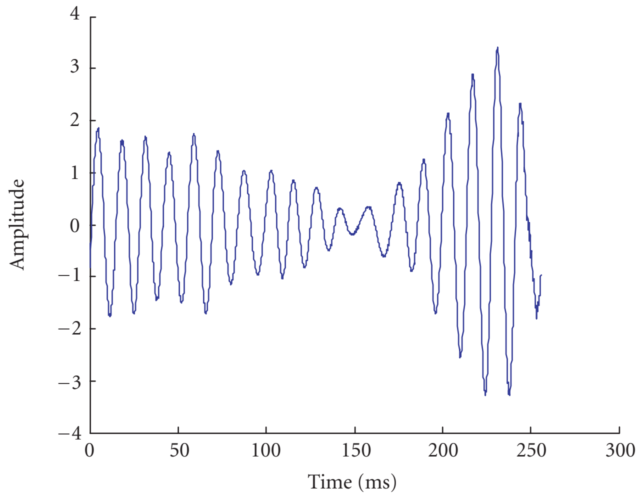



Figure 4: Direct inverse impulse response.

minimum-phase impulse response $g_{mp}(n)$ to be of a very long duration (Figure 7).

When using the standard version for a significant value of $L$, ($L = 2$), (Figure 6) we observed that all the poles of $G_{mp}(k)$ were pushed together towards the origin of the unit circle too, resulting in an inverse minimum-phase impulse response $g_{mp}(n)$ (Figure 7) to be reduced in time too.

When using a modified version, in order to gradually reduce the length of the inverse minimum-phase impulse response $g_{mp}(n)$, only the most dominant pole needed to be replaced by a new pole with smaller radius. This pole corresponded to a $Q/l$ factor with the same value of $l$, ($l = L = 2$), but at the same frequency. In Figure 5 we can see the inverse minimum-phase frequency response of $\widetilde{G}_{mp}(k)$, where only the most dominant pole appears to be pushed towards the origin, with $l = 2$.

Figure 7 also shows the corresponding inverse minimum-phase impulse response of $\widetilde{g}_{mp}(n)$. Interestingly, its duration is not reduced here too. This may correspond to a desired magnitude equalization (Figure 8), if the system impulse response was minimum phase (no phase distortion effects). This is because the magnitude spectrum of the second case (modified method, $l = 2$, $\Delta$(dB) = 0.7) is flatter than that of the first case (standard method, $L = 2$, $\Delta$(dB) = 2.4). This means less magnitude distortion of the system.

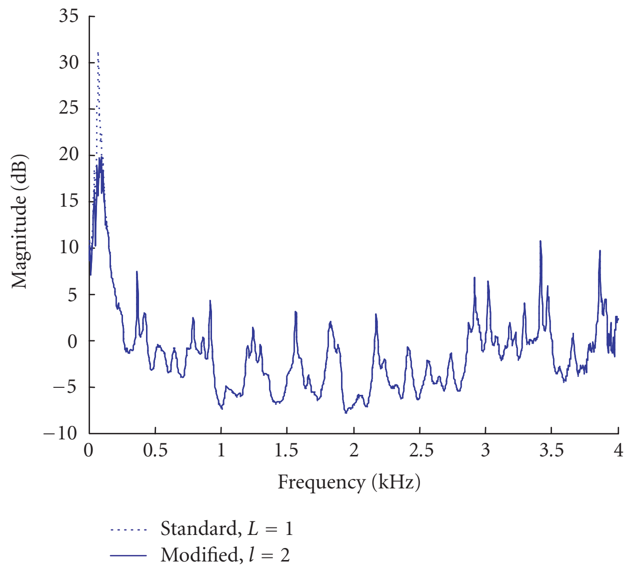



Figure 5: Inverse minimum-phase frequency response $G_{mp}(k)$ calculated by the different methods.

### 4.3. *Performance testing*

The experiment was performed by developing models in Matlab and Simulink and carrying out listening tests using headphones. A reproduced speech signal of few seconds in duration was generated by filtering a clean speech (male and female measured in anechoic chamber) by the measured impulse response of the car interior (Figure 3). In order to avoid undesirable convolution effects, we considered a sufficient large number $N = 8192$ for DFT computations. The reproduced speech signal was then filtered using equalizing filters calculated by the standard method with $L = 1$ and $L = 2$ (Figure 7) and the modified version ($P = 1$), respectively. For the latter case, the inverse impulse responses corresponded to each error criterion, function of the parameter $l$ such as that of Figure 7 with $l = 2$, for example. Test signals were played to ten untrained listeners with normal hearing at a comfortable listening level. The qualitative assessment of the test signals was based on subjective judgment of three listening sessions per each recording scheduled on six consecutive days.

The first signal was always chosen to be clean speech, while the reproduced unequalized and partially equalized speech signals were played in random order. The reproduced speech signals corresponded to the objective error criteria of 5 dB (unequalized signal for $l = 0$), 0 dB (magnitude equalized signal for $l = 1$), and $0.3 \le \Delta(\text{dB}) \le 3$ (partially equalized

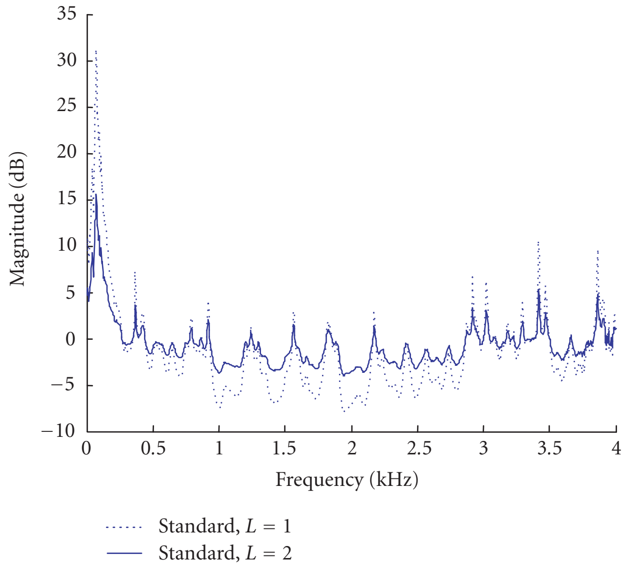



FIGURE 6: Inverse minimum-phase frequency response $G_{\text{mp}}(k)$ calculated by the different methods.

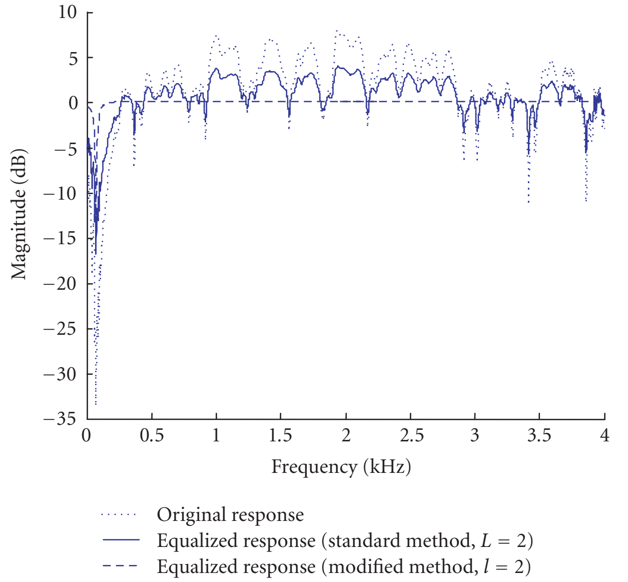



FIGURE 8: Magnitude response equalization.

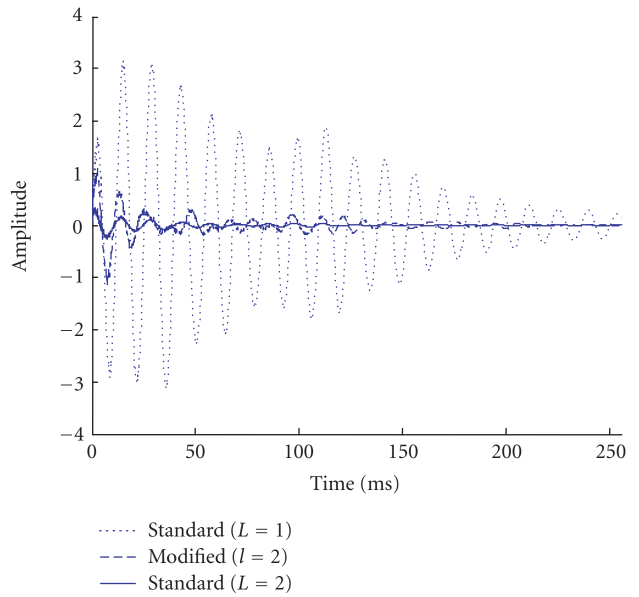



FIGURE 7: Inverse minimum-phase impulse responses $g_{\text{mp}}(n)$ calculated by the different methods.

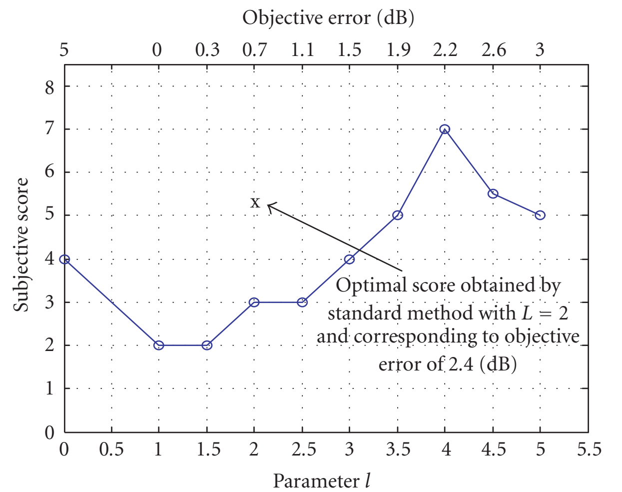



FIGURE 9: Subjective scores of the sound quality as a function of the parameter $l$. Each circle represents the average of 180 observations: 18 for each 10 listeners.

signals for $l > 1$), respectively. The quantification of subjective judgments was performed according to the following scale [4]:

(i) 7, 8: good;
(ii) 5, 6: fair;
(iii) 3, 4: poor;
(iv) 1, 2: bad.

Number 8 denotes a sound quality equivalent to the clean speech. The final result was calculated as a mean of the individual listening results (18 each) for each of 10 subjects. The results are shown in Figure 9 as a function of parameter $l$, ranging from $l = 0$ (unequalized signal) to $l = 5$ (partially equalized signal).

The results confirmed those reported in [4] with higher accuracy. The highest score corresponds to the optimal quality of speech. That means no perception of phase distortion (like a bell chime sounded at the background, when $l < 3$), no echo and less magnitude distortion caused by the system.

The results also show the sensitivity of the listener's ear to small gradual response variations (a variation of less than

0.5 dB of objective error corresponds to a significant variation of subjective score of the sound quality); although the participants in the experiment were nonexpert listeners.

## 5. CONCLUSIONS

In this paper a modified version of the standard homomorphic method for minimum-phase inverse filter design for non-minimum-phase impulse responses equalization is presented. This version is useful in cases of partial magnitude equalization, where the dominant zeros density of the system is not very high. Although it is used in this work as an additional optimizing tool for the psychoacoustic quality measurement of speech, this alternative approach is advantageous in case of the direct inverse filtering (minimum-phase system) when perfect equalization of a small reverberant room is not desired.

## APPENDIX

Proof for the relation (13).

The real part of the complex cepstrum of $h(n)$ in (6) is defined as the inverse DFT of the function

$$\hat{H}(k) = \log |H(k)|. \tag{A.1}$$

Applying the direct DFT on the real cepstrum of the minimum-phase $h_{\text{mp}}(n)$ in the relation (7) for $L = 1$ leads to

$$\hat{H}_{\text{mp}}(k) = \log |H_{\text{mp}}(k)|. \tag{A.2}$$

For $L \neq 1$, the relation (A.2) becomes

$$\hat{H}_{\text{mp}}(k) = \frac{1}{L}\log |H_{\text{mp}}(k)|. \tag{A.3}$$

Using the relation (3), the minimum-phase part $H_{\text{mp}}(k)$ of the relation (9) can be expressed as follows:

$$H_{\text{mp}}(k) = \exp\left(\frac{1}{L}\log |H_{\text{mp}}(k)|\right) = \exp\left(\frac{1}{L}\log |H(k)|\right). \tag{A.4}$$

Therefore, the inverse of $H_{\text{mp}}(k)$, $G_{\text{mp}}(k)$ is given by

$$G_{\text{mp}}(k) = \exp\left(-\frac{1}{L}\log |H(k)|\right), \tag{A.5}$$

or

$$\log |G_{\text{mp}}(k)| = -\frac{1}{L}\log |H(k)|. \tag{A.6}$$

## ACKNOWLEDGMENT

The authors would like to thank the reviewers for their helpful comments and suggestions.

## REFERENCES


[1] S. J. Elliott and P. A. Nelson, "Multiple-point equalization in a room using adaptive digital filters," *Journal of the Audio Engineering Society*, vol. 37, no. 11, pp. 899–907, 1989.

[2] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1975.

[3] J. N. Mourjopoulos, "Digital equalization of room acoustics," *Journal of the Audio Engineering Society*, vol. 42, no. 11, pp. 884–900, 1994.

[4] B. D. Radlovic and R. A. Kennedy, "Nonminimum-phase equalization and its subjective importance in room acoustics," *IEEE Transactions on Speech and Audio Processing*, vol. 8, no. 6, pp. 728–737, 2000.

[5] L. D. Fielder, "Analysis of traditional and reverberation-reducing methods of room equalization," *Journal of the Audio Engineering Society*, vol. 51, no. 1/2, pp. 3–26, 2003.

[6] A. V. Oppenheim and R. W. Schafer, *Discrete Time Signal Processing*, Prentice Hall, Upper Saddle River, NJ, USA, 1989.

[7] S. T. Neely and J. B. Allen, "Invertibility of a room impulse response," *The Journal of the Acoustical Society of America*, vol. 66, no. 1, pp. 165–169, 1979.

[8] B. D. Radlovic and R. A. Kennedy, "Iterative cepstrum-based approach for speech dereverberation," in *Proceedings of the 5th International Symposium on Signal Processing and Its Applications (ISSPA '99)*, vol. 1, pp. 55–58, Brisbane, Australia, August 1999.

[9] D. Preis, "Phase distortion and phase equalization in audio signal processing—a tutorial review," *Journal of the Audio Engineering Society*, vol. 30, no. 11, pp. 774–794, 1982.

[10] A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Välimäki, "Modal equalization of loudspeaker-room responses at low frequencies," *Journal of the Audio Engineering Society*, vol. 51, no. 5, pp. 324–343, 2003.

[11] M. Karjalainen, P. A. A. Esquef, P. Antsalo, A. Mäkivirta, and V. Välimäki, "Frequency-zooming ARMA modelling of resonant and reverberant systems," *Journal of the Audio Engineering Society*, vol. 50, no. 12, pp. 1012–1029, 2002.

[12] M. Bellanger, *Traitement Numérique du Signal*, Dunod, Paris, France, 1998.

[13] Y. Haneda, S. Makino, and Y. Kaneda, "Common acoustical pole and zero modeling of room transfer functions," *IEEE Transactions on Speech and Audio Processing*, vol. 2, no. 2, pp. 320–328, 1994.

[14] M. Kunt, *Traitement Numérique des Signaux*, Dunod, Paris, France, 1981.

[15] J. R. Hopgood and P. J. W. Rayner, "Blind single channel deconvolution using nonstationary signal processing," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 5, pp. 476–488, 2003.


**Ahfir Maamar** received his "Ingeniorat" in 1990 and "Magister" in 1997 in electronics, both from the University of Blida (Algeria). He is currently a Lecturer at University of Laghouat (Algeria) and Visiting Researcher to Applied DSP and VLSI Systems Laboratory of the University of Westminster, London, UK. His areas of interest include room acoustics, acoustic/audio/ electroacoustic signal processing, and inverse filtering.

**Izzet Kale** holds the B.S. (honors) degree in electrical and electronic engineering from the Polytechnic of Central London (England), the M.S. degree in the design and manufacture of microelectronic systems from Edinburgh University (Scotland), and the Ph.D. degree in techniques for reducing digital filter complexity from the University of Westminster (England). He joined the staff of the University of Westminster (formerly the Polytechnic of Central London) in 1984 and he has been with them since. He is currently a Professor of applied DSP and VLSI systems, leading the Applied DSP and VLSI Research Group at the University of Westminster. His research and teaching activities include digital and analogue signal processing, silicon circuit and system design, digital filter design and implementation, A/D and D/A sigma-delta converters. He is currently working on efficiently implementable, low-power DSP algorithms/architectures and sigma-delta modulator structures for use in the communications and biomedical industries.

**Artur Krukowski** holds the Ph.D. degree in DSP from the University of Westminster (London in 1999), the M.S. degree in instrumentation and measurement (I & M) from the Warsaw University of Technology (Warsaw in 1992), and the M.S. degree in DSP from the University of Westminster (London in 1993). In 1993 he joined the staff of the University of Westminster. His main areas of interest include multirate digital signal processing for telecommunication systems, efficient low-level implementation of multirate fixed/floating-point digital filters, digital audio and video broadcasting, e-teaching and e-Learning technologies. From 2004 he is associated also with the National Research Center "Demokritos" in Athens (Greece) where he carries out advanced research in indoor/outdoor positioning technologies as enabling technologies for the provision of advanced location-based services.

**Berkani Daoud** received the Engineer Diploma and Master's degree with Red Award from Polytechnic Institute of Kiev in 1977, then the Magister and Sc.D. degrees from the Ecole Nationale Polytechnique (ENP), Algiers. In 1979, he became a Lecturer, Associated Professor, then full Professor teaching signal processing and information theory in the Department of Electronics of ENP. During this period, his research activities involved the applications of signal processing and the source coding theory. In 1992, he joined the Department of Electrical Engineering of University of Sherbrooke, Canada, where he taught signal processing. He was a member of the Speech Coding team of the University of Sherbrooke. He has been conducting research in the area of speech coding and speech processing in adverse conditions. Since 1994, he backs to the Department of Electrical and Computer Engineering of ENP. His current research interests include signal and communications, information theory concepts, and clustering algorithm, applied to speech and image processing.

# ROOM EQUALIZATION BASED ON ITERATIVE SIMPLE COMPLEX SMOOTHING OF ACOUSTIC IMPULSE RESPONSES

*Ahfir Maamar [(1)], Izzet Kale [(2)], Arthur Krukowski [(3)] and Berkani Daoud [(4)]*

[(1)] Department of Informatics, University of Laghouat, BP: 37G, Laghouat, ALGERIA, Email: m.ahfir@mail.lagh-univ.dz.
[(2)], [(3)] Applied DSP and VLSI Research Group, Department of Electronic Systems, University of Westminster, 115 New Cavendish Street, London, UK. Email: kalei@westminster.ac.uk,
[(4)] Department of Electronics, ENPA, BP: 182, Alger, ALGERIA. Email: dberkani@hotmail.com.

**ABSTRACT**

This paper presents a room equalization method based on iterative simple complex smoothing of measured acoustic impulse responses. This is useful in cases of long duration impulse responses. Corresponding time reduced impulse responses are derived which conform to perceptual principles. The smoothed impulse responses are then used to design equalization filters. Results from an audio-conferencing reverberant room using objective and subjective tests show that we can improve the measured and perceived quality of audio reproduction.

## 1. INTRODUCTION

Direct equalization of room acoustics based on inverse filtering of measured mixed-phase impulse responses introduces a number of theoretical and practical problems (ideal digital equalization). Particularly, the required extremely long lengths of the inverse filters (magnitude and phase equalization) and their sensitivity to possible changes in the listener position inside the room [1], [2]. When the smoothed room impulse response was used for the design of `inverse filters` for audio or acoustic digital equalization applications, it was found that the effect of smoothing was desirable since it was allowing the design of filters with lower sensitivity – from the perceptual point of view - to possible changes in the listener position inside the room [3], [4], [5]. The work presented here is based on the concept of complex smoothing performed by the simple case of constant bandwidth filters for `inverse filter` design for acoustic impulse responses equalization. This concept is adopted because modified impulse response functions could be derived, which would present functions of reduced complexity and also be in agreement with the perceptual principles. The work presented in this paper is organized as follows: Section 2 presents the theory of the simple complex smoothing [3]. In Section 3 an equalization filter design method is proposed, which is based on iterative simple complex smoothing. Section 4 shows the magnitude equalization performance results achieved using this proposed method in the context of a real audio-conferencing room (cnet-LANNION-FRANCE).

## 2. SIMPLE COMPLEX SMOOTHING

Let us consider a discrete-time room impulse response $h(n)$ and its discrete-frequency response $H(k)$, where $0 \le n \le N-1$, $0 \le k \le N-1$ and $N$ is the number of samples, representing the impulse response length. Then the complex smoothing operation for the simple case of constant bandwidth filter can be described as a circular convolution [3]:

$$H_{cs}(k) = H(k) \otimes W_{sm}(k) \tag{1}$$

Where $\otimes$ denotes the operation of circular convolution and $W_{sm}(k)$ is a spectral smoothing function having the general form of a low-pass filter.

This is equivalent in the time domain to:

$$h_{cs}(n) = h(n) N w_{sm}(n) \tag{2}$$

where $w_{sm}(n)$ being the inverse DFTsequence of $W_{sm}(k)$.

To consider this spectral smoothing function in a half-window sense for both parts of the symmetric spectrum $W_{sm}(k)$ must be written as:

$$W_{sm}(k) = \begin{cases} \dfrac{b-(b-1)\cos(\pi k/m)}{2b(m+1)-1}, & k = 0,1,\dots,m \\ \dfrac{b-(b-1)\cos(\pi(k-N)/m)}{2b(m+1)-1}, & k = N-m, N-(m-1),\dots, N-1 \\ 0 & k = m+1,\dots, N-(m+1) \end{cases} \tag{3}$$

Where $m$ (samples) is defined as the smoothing index corresponding to the length of the half window. When $b=1$, this function represents an ideal low-pass filter (rectangular frequency smoothing function),

$$W_{sm}(k) = \begin{cases} \dfrac{1}{2m+1}, & k \in \{0,1,\dots,m\} \cup \{N-m, N-(m-1),\dots, N-1\} \\ 0 & k \in \{m+1,\dots, N-(m+1)\} \end{cases} \tag{4}$$

In this case the corresponding time window function which is represented with a *sinc* function $w_{sm}(n)$ can be evaluated easily as:

$$w_{sm}(n)=\frac{1}{N}\frac{\sin(\pi n(2m+1)/N)}{\pi n(2m+1)/N} \quad (5)$$

Given that $H(k)$ is a complex function; in general $W_{sm}(k)$ should be a complex function. However, the preceding expressions represent it as a real function, assuming it to be a zero-phase function. This assumption was adopted because of physical considerations, since with smoothing it is required to avoid imposition of any unwanted effects on the phase of the original function and it is also desirable to maintain the half-window time profile appropriate for capturing transient data of the acoustic impulse response that are significant to the listener (direct signal and first reflections).

## 3. EQUALIZATION FILTER DESIGN BASED ON ITERATIVE SIMPLE COMPLEX SMOOTHING

In order to overcome the known problems of direct equalization as stated in the introduction, we propose in this section an equalization method based on iterative simple complex smoothing of room impulse responses. In this method the original room impulse response $h(n)$ is assumed to be the result of smoothing operation in the time domain of an initial one (its length was $2N$ ) by a half smoothing window $w_{sm}(n)$ of $N$ samples. Then in the first iteration $h(n)$ will be halved in length after the application of a half smoothing window $w_{sm}(n)$ of $\frac{N}{2}$ samples. As a result a reduced-length complex smoothed impulse response $h_{cs}(n)$ will be produced and a minimum-phase inverse impulse response $g_{cs,mp}(n)$ for smoothed response magnitude equalization will be evaluated by using the homomorphic method [6], since the smoothed response may present a mixed-phase character. This minimum-phase inverse impulse response $g_{cs,mp}(n)$ will be the `equalization filter`, employed on the original acoustic impulse response $h(n)$ according to the following expression:

$$h_{eq}(n)=h(n)\otimes g_{cs,mp}(n) \quad (6)$$

The iteration process is repeated only if the result of the smoothing operation will be meaningful, that is, the early time components (direct signal and first reflections) of the already smoothed impulse response $h_{cs}(n)$ must be preserved when halving it again with a half smoothing window $w_{sm}(n)$. The procedure may be stated as follows:

1. Set the iteration index $i$ to 0 and choose the parameters $b$ and $m$.
2. Compute the spectral smoothing window $W_{sm}(k)$ using (3) for $k=0,1,2,....,\frac{N}{2^i}$.
3. Compute the corresponding time smoothing window $w_{sm}(n)$ by using the IDFT and apply it for $n=0,1,2,....,\frac{N}{2^{i+1}}$.
4. Compute the smoothed impulse response $h_{sm}{}^{i+1}(n)$ as:

$$h_{cs}{}^{i+1}(n)=h_{cs}{}^{i}(n)\frac{N}{2^i}w_{sm}(n) \quad (7)$$

where $h_{cs}{}^{0}(n)=h(n)$.

5. Evaluate the corresponding minimum-phase inverse impulse response $g_{cs,mp}{}^{i+1}(n)$ using the homomorphic method
6. Compute the equalized room impulse response $h_{eq}(n)$ as:

$$h_{eq}(n)=h(n)\otimes g_{cs,mp}{}^{i+1}(n) \quad (8)$$

7. Increment $i$, $i=i+1$.

## 4. RESULTS

The proposed method was applied to a measured acoustic impulse response of an audio-conferencing room at a sampling frequency of 16 kHz and recorded to $N=8192$ (Figure 1) [8]. In this example three iterations ( $i=0,1,2$ ) of the smoothing process using an initial half smoothing window (Figure 2 - $b=0.5$ and $m=3$ ) was needed to preserve the initial transient portion of the acoustic impulse response. A reduced-length of the smoothed response $h_{cs}{}^{3}(n)$ (Figure 3) was produced and then used to construct the appropriate minimum-phase inverse impulse response $g_{cs,mp}{}^{3}(n)$ by using an iterative version of the homomorphic method [7] with an optimum number of iterations $L=8$. This means that $g_{cs,mp}{}^{3}(n)$ is constructed by 8 successive convolutions of partially calculated minimum-phase inverse impulse responses $g_{cs,mp}{}^{1/8}(n)$ (Figure 4). The particular choice of this iterative version over the standard method is because of reduced time domain aliasing given a fixed number $N$ for DFT computation.

In order to assess the performance of our algorithm, we used an objective frequency domain error criterion, which estimates the standard deviation of the magnitude response from a constant level [2]. The error criterion $\Delta(dB)$ is given as follows:

$$\Delta=\left[\frac{1}{N}\sum_{k=0}^{N-1}\left(10\log_{10}|H(k)|-H_m\right)^2\right]^{1/2} \quad (9)$$

where:

$$H_m=\frac{1}{N}\sum_{k=0}^{N-1}10\log_{10}|H(k)| \quad (10)$$

and evaluated for both the original and equalized responses. This error criterion was also evaluated for both a second acoustic impulse response which is different to the initial one, and its corresponding equalized response using the same inverse filter used to equalize the initial one; assessing in this way potential robustness of the proposed method to possible changes in the listener position inside the room.

The results for magnitude equalization are tabulated below in the form of the objective error criterion evaluated for both original and equalized responses. The magnitude responses of the two cases (initial and second) for both original and equalized are shown in Figure 5.

**Table 1.** Objective error criterion (dB) evaluated for both original and equalized responses.

| Acoustic Response | Original Response (dB) | Equalized Response (dB) |
|---|---|---|
| Initial | 8.31 | 3.54 |
| Second | 7.46 | 5.26 |

These results indicate a significant reduction of the objective error criterion as it can be seen as spectral magnitude deviation from flatness for both cases (Figure 5). The subjective testing was also evaluated by listening tests through headphones using Matlab`s Audio Tools and a reverberated anechoic speech as the original signal test. The subjective results indicated a significant perceptual preference for the equalized signal over the original. These results can be considered as promising for the case of possible changes in the listener position inside the room.

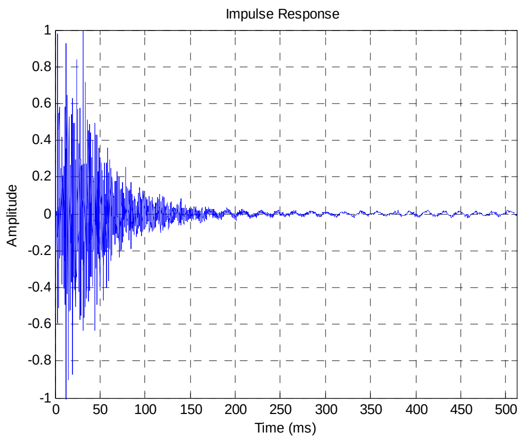



**Figure 1.** Measured acoustic impulse response

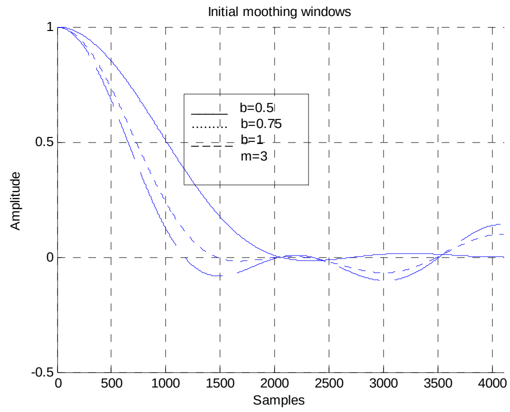



**Figure 2.** Initial smoothing windows for $i = 0$

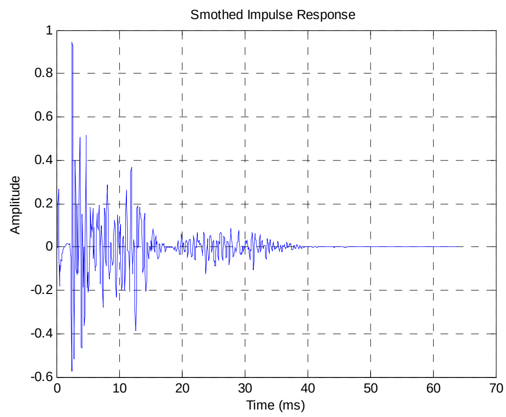



**Figure 3.** Smoothed impulse response $h_{cs}^{3}(n)$

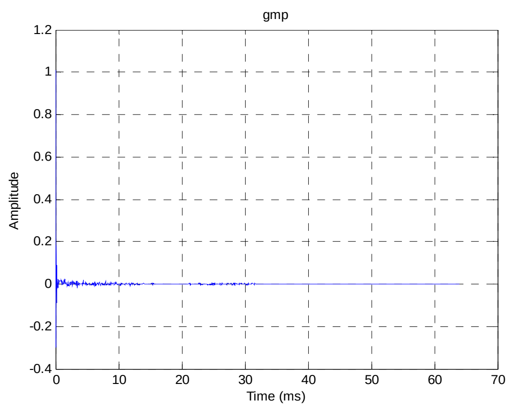
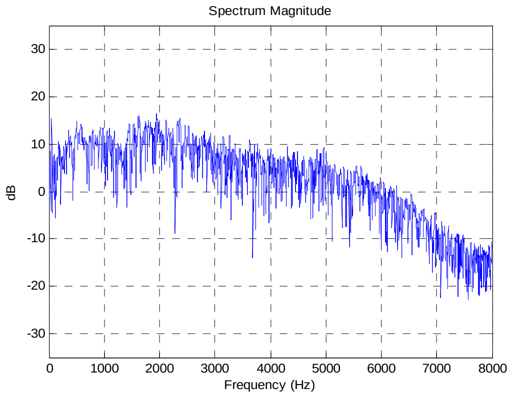
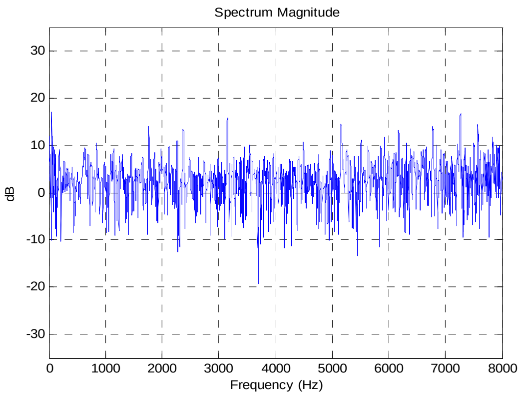



**Figure 4.** Partial minimum-phase inverse impulse response $g_{cs,\,mp}^{1/8}(n)$ calculated using the iterative version of the homomorphic method with $L = 8$ .

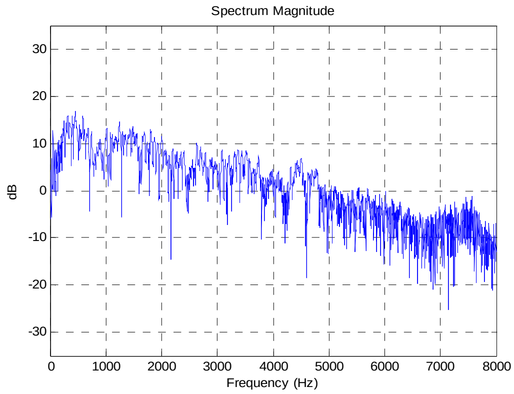



(a)

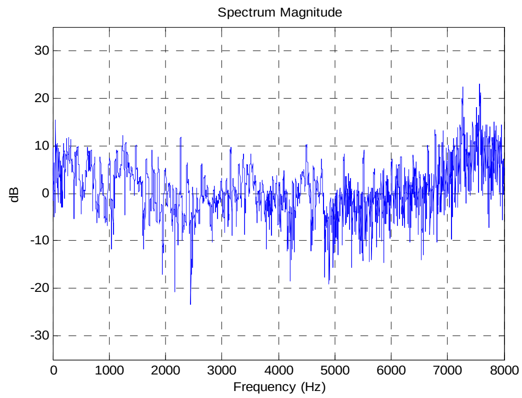



(b)

**Figure 5 – 1.** Deviation of magnitude response from flatness
(a) Initial original response, b) Equalized response




(a)




(b)

**Figure 5 – 2.** Deviation of magnitude response from flatness
a) Second original response, b) Equalized response

## 5. CONCLUSIONS

Equalization filters design based on iterative simple complex smoothing has been proposed for room equalization. The algorithm can be useful in cases of long duration acoustic impulse responses since it results in impulse responses of reduced complexity which preserve the initial transient data that are significant to the listener. The algorithm also gives the possibility of stopping the smoothing process when the corresponding equalization filter produces some pre-determined equalized magnitude response.

## 6. REFERENCES

[1] B. D. Radlovic and R. A. Kennedy, “Non-minimum phase equalization and its subjective importance in room acoustics”, IEEE Transactions on Speech and Audio Processing, vol. 8, no. 6, Nov 2000.

[2] J. N. Mourjopoulos, "Digital Equalization of Room Acoustics", J. Audio. Eng. Soc, Vol 42, N°11, 1994 November.

[3] P.Hatziantonio and J.Mourjoupoulos, " Generalized Fractional-Octave Smoothing of Audio and Acoustic Responses", J. Audio. Eng. Soc, vol 48, No 4, pp 259-280, 2000 April.

[4] P. M .Clarkson, J.Mourjopoulos and J.K.Hammond, "Spectral, Phase and Transient Equalization for Audio Systems", J. Audio. Eng. Soc, vol 33, pp 127-132, 1985 Mar.

[5] S. Salamouris, K. Politopoulos, V. Tsakiris and J. Mourjopoulos, “Digital System for Loudspeaker and Room Equalization ", presented at the 98$^{th}$ Convention of the Audio Engineering Society, J. Audio. Eng. Soc, (Abstracts), vol 43, pp 396, 1985 May.

[6] Neely S. T. and J. B. Allen, “Invertibility of a room impulse response”, J. Acoust. Soc. Am, vol. 66, no. 1, pp. 165-169, July 1979.

[7] B. D. Radlovic and R. A. Kennedy, “Iterative cepstrum-based approach for speech dereverberation ", ISSPA` 99, Brisbane, Australia, 22-25 August 1999.

[8] J.P.Jullien, A.Gilloire and A.Saliou, “Mesure des reponses impulsionnelles en acoustique”, Note technique NT/LAA/TSS/181, Novembre 1984.

# Implementation of a Reduced Complexity FIR Inverse Filter on DSP Board for Room Equalization

Ahfir Maamar [(1)], Izzet Kale [(2)], Berkani Daoud [(3)] and Ediz Cetin [(2)]
[(1)] Department of Informatics, University of Laghouat, BP: 37G, Laghouat, ALGERIA, Email: m.ahfir@mail.lagh-univ.dz.
[(2)] Applied DSP and VLSI Research Group, Department of Electronic Systems, University of Westminster, 115 New Cavendish Street, London, UK. Email: kalei@westminster.ac.uk, E.Cetin@westminster.ac.uk
[(3)] Department of Electronics, ENPA, BP: 182, Alger, ALGERIA. Email: dberkani@hotmail.com.

*Abstract* - **In this paper, a technique for dereverberation systems design based on partial equalization for room acoustics is presented. In this technique, an iterative simple smoothing process is applied to the original impulse response of a large room before its inversion. Corresponding time reduced impulse responses are derived which conform to perceptual principles. The smoothed impulse responses are then used to design equalization filters using the Homomorphic method. Equalization results for a reverberant audio-conferencing room have been validated by listening tests after implementation of a reduced complexity FIR inverse filter on Texas Instruments DSP board.**

## I. INTRODUCTION

Reverberation phenomenon is caused by the room effect where the useful source is placed. This is because of acoustic wave reflections. Hands-free or/and loud-speaker systems are very sensitive to reverberation in communications and, in general, when a sound recording is taken in an enclosure without a particular acoustical treatment. When reverberation is relatively significant, the source seems to be subjectively far away and then the speech intelligibility is degraded.

Dereverberation is to make the desired signal completely or partially clean of the room effect. A perfect dereverberation would leave only the wave coming directly from the source and would produce the same effect than a sound recording taken in an anechoic chamber. This is not required in room acoustics because a moderate room effect can make the speech to sound natural.

A first approach of dereverberation is an inversion of the acoustic channel represented by its measured impulse response, which provides the room effect between the source and the microphone installed in the room. This approach is not realistic because the frequency response of the room has many zeros which rapidly vary with the microphone position in the room. Therefore, it is required to use methods taking into account the temporal and spatial aspects of the reverberation phenomenon [1].

Some methods based on direct inversion of the acoustic channel have been proposed by some authors [2], [3], [4], but they still need modifications to be adapted to corresponding contexts, to better respond to the desired speech quality. In this paper, we present an equalization technique (inversion method) of the acoustic channel, recommended in the context of a large room, such as audio-conferencing and video-conferencing rooms.

The goal is to design stable equalizers Finite Impulse Responses (FIR) digital filters (FIR inverse filters) taking into account temporal and spatial aspects of the reverberation phenomenon.

This technique which is based on the application of an iterative simple smoothing process to the original acoustic impulse response before its inversion, allows reducing simultaneously the rebeverberation effect and the complexity of the designed FIR inverse filters for implementation [5].

The work presented in this paper is organized as follows: Section 2 presents the theory of the simple complex smoothing [6]. In Section 3 an equalization filter design method is proposed, which is based on iterative simple complex smoothing. Section 4 shows the magnitude equalization performance results achieved using this proposed method in the context of a real audio-conferencing room (CNET-LANNION-FRANCE). Section 5 discusses the listening tests results obtained for a large audio-conferencing room, after implementation of an equalization filter example on Texas Instruments DSP board.

## II. SIMPLE COMPLEX SMOOTHING

Direct equalization of room acoustics based on inverse filtering of measured mixed-phase impulse responses introduces a number of theoretical and practical problems (ideal digital equalization). Particularly, the required extremely long lengths of the inverse filters (magnitude and phase equalization) and their sensitivity to possible changes in the listener position inside the room [3], [7]. When the smoothed room impulse response was used for the design of `inverse filters` for audio or acoustic digital equalization applications, it was found that the effect of smoothing was desirable since it was allowing the design of filters with lower sensitivity – from the perceptual point of view - to possible changes in the listener position inside the room [8], [9], [10]. The work presented here is based on the concept of complex smoothing performed by the simple case of constant bandwidth filters for `inverse filter` design for acoustic impulse responses equalization. This concept is adopted because modified impulse response functions could be derived, which would present functions of reduced complexity and also be in agreement with the perceptual principles.

Let us consider a discrete-time room impulse response $h(n)$ and its discrete-frequency response $H(k)$, where $0 \le n \le N-1$, $0 \le k \le N-1$ and $N$ is the number of samples, representing the impulse response length. Then the complex smoothing operation for the simple case of constant bandwidth filter can be described as a circular convolution [8]:

$$H_{cs}(k) = H(k) \otimes W_{sm}(k) \quad (1)$$

Where $\otimes$ denotes the operation of circular convolution and $W_{sm}(k)$ is a spectral smoothing function having the general form of a low-pass filter.
This is equivalent in the time domain to:

$$h_{cs}(n) = h(n) N w_{sm}(n) \quad (2)$$

where $w_{sm}(n)$ being the inverse DFT sequence of $W_{sm}(k)$.

To consider this spectral smoothing function in a half-window sense for both parts of the symmetric spectrum $W_{sm}(k)$ must be written as:

$$W_{sm}(k) = \begin{cases} \dfrac{b-(b-1)\cos(pk/m)}{2b(m+1)-1}, & k = 0,1,\ldots,m \\ \dfrac{b-(b-1)\cos(p(k-N)/m)}{2b(m+1)-1}, & k = N-m, N-(m-1),\ldots, N-1 \\ 0 & k = m+1,\ldots, N-(m+1) \end{cases} \quad (3)$$

Where $m$ (samples) is defined as the smoothing index corresponding to the length of the half window. When $b=1$, this function represents an ideal low-pass filter (rectangular frequency smoothing function),

$$W_{sm}(k) = \begin{cases} \dfrac{1}{2m+1}, & k \in \{0,1,\ldots,m\} \cup \{N-m, N-(m-1),\ldots, N-1\} \\ 0 & k \in \{m+1,\ldots, N-(m+1)\} \end{cases} \quad (4)$$

In this case the corresponding time window function which is represented with a *sinc* function $w_{sm}(n)$ can be evaluated easily as:

$$w_{sm}(n) = \frac{1}{N} \frac{\sin(pn(2m+1)/N)}{pn(2m+1)/N} \quad (5)$$

Given that $H(k)$ is a complex function; in general $W_{sm}(k)$ should be a complex function. However, the preceding expressions represent it as a real function, assuming it to be a zero-phase function. This assumption was adopted because of physical considerations, since with smoothing it is required to avoid imposition of any unwanted effects on the phase of the original function and it is also desirable to maintain the half-window time profile appropriate for capturing transient data of the acoustic impulse response that are significant to the listener (direct signal and first reflections).

## III. EQUALIZATION FILTER DESIGN BASED ON ITERATIVE SIMPLE COMPLEX SMOOTHING

In order to overcome the known problems of direct equalization as stated in the previous section, we propose in this section an equalization method based on iterative simple complex smoothing of room impulse responses. In this method the original room impulse response $h(n)$ is assumed to be the result of a smoothing operation in the time domain of an initial one (its length was $2N$) by a half smoothing window $w_{sm}(n)$ of $N$ samples. Then in the first iteration $h(n)$ will be halved in length after the application of a half smoothing window $w_{sm}(n)$ of $\frac{N}{2}$ samples. As a result a reduced-length complex smoothed impulse response $h_{cs}(n)$ will be produced and a minimum-phase inverse impulse response $g_{cs,mp}(n)$ for smoothed response magnitude equalization will be evaluated by using the homomorphic method [2], since the smoothed response may present a mixed-phase character. This minimum-phase inverse impulse response $g_{cs,mp}(n)$ will be the `equalization filter`, employed on the original acoustic impulse response $h(n)$ according to the following expression:

$$h_{eq}(n) = h(n) \otimes g_{cs,mp}(n) \quad (6)$$

The iteration process is repeated only if the result of the smoothing operation will be meaningful, that is, the early time components (direct signal and first reflections) of the already smoothed impulse response $h_{cs}(n)$ must be preserved when halving it again with a half smoothing window $w_{sm}(n)$. The procedure may be stated as follows:

1. Set the iteration index $i$ to 0 and choose the parameters $b$ and $m$.
2. Compute the spectral smoothing window $W_{sm}(k)$ using (3) for $k = 0,1,2,\ldots,\frac{N}{2^i}$.

3. Compute the corresponding time smoothing window $w_{sm}(n)$ by using the IDFT and apply it for $n = 0,1,2,....,\frac{N}{2^{i+1}}$.
4. Compute the smoothed impulse response $h_{sm}^{i+1}(n)$ as:

$$h_{cs}^{i+1}(n) = h_{cs}^{i}(n)\frac{N}{2^i}w_{sm}(n) \qquad (7)$$

where $h_{cs}^{0}(n) = h(n)$.

5. Evaluate the corresponding minimum-phase inverse impulse response $g_{cs,mp}^{i+1}(n)$ using the Homomorphic method.
6. Compute the equalized room impulse response $h_{eq}(n)$ as:

$$h_{eq}(n) = h(n) \otimes g_{cs,mp}^{i+1}(n). \qquad (8)$$

7. Increment $i$, $i = i+1$ and go to step 2.

## IV. Simulation Results

The proposed method was applied to a measured acoustic impulse response of an audio-conferencing room at a sampling frequency of 16 kHz and recorded to $N = 8192$ (Figure 1) [11]. In this example three iterations ($i = 0,1,2$) of the smoothing process using an initial half smoothing window (Figure 2 - $b = 0.5$ and $m = 3$) was needed to preserve only the initial transient portion of the acoustic impulse response. A reduced-length of the smoothed response $h_{cs}^{3}(n)$ (Figure 3) was produced and then used to construct the appropriate minimum-phase inverse impulse response $g_{cs,mp}^{3}(n)$ by using an iterative version of the homomorphic method [12] with an optimum number of iterations $L = 8$. This means that $g_{cs,mp}^{3}(n)$ is constructed by 8 successive convolutions of partially calculated minimum-phase inverse impulse responses $g_{cs,mp}^{1/8}(n)$ (Figure 4). The particular choice of this iterative version over the standard method is because of reduced time domain aliasing given a fixed number $N$ for DFT computation.

In order to assess the performance of our algorithm, we used an objective frequency domain error criterion, which estimates the standard deviation of the magnitude response from a constant level [7]. The error criterion $\Delta(dB)$ is given as follows:

$$\Delta = \left[\frac{1}{N}\sum_{k=0}^{N-1}\left(10\log_{10}|H(k)| - H_m\right)^2\right]^{1/2} \qquad (9)$$

where:

$$H_m = \frac{1}{N}\sum_{k=0}^{N-1}10\log_{10}|H(k)| \qquad (10)$$

and evaluated for both the original and equalized responses. This error criterion was also evaluated for both a second acoustic impulse response which is different to the initial one, and its corresponding equalized response using the same inverse filter used to equalize the initial one; assessing in this way potential robustness of the proposed method to possible changes in the listener position inside the room.

The results for magnitude equalization are tabulated below in the form of the objective error criterion evaluated for both original and equalized responses. The magnitude responses of the two cases (initial and second) for both original and equalized are shown in Figure 5.

TABLE 1
Objective error criterion (dB) evaluated for both original and equalized responses.

| Acoustic Response | Original Response (dB) | Equalized Response (dB) |
|---|---|---|
| Initial | 8.31 | 3.54 |
| Second | 7.46 | 5.26 |

These results indicate a significant reduction of the objective error criterion as it can be seen as spectral magnitude deviation from flatness for both cases (Figure 5). The subjective testing was also evaluated by listening tests through a headphone using Matlab`s Audio Tools and a reverberated anechoic speech as the original signal test. The subjective results indicated a significant perceptual preference for the equalized signal over the original. These results can be considered as promising for the case of possible changes in the listener position inside the room.

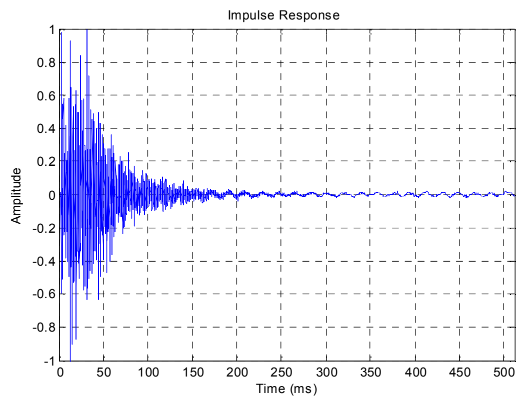



Fig 1 Measured acoustic impulse response

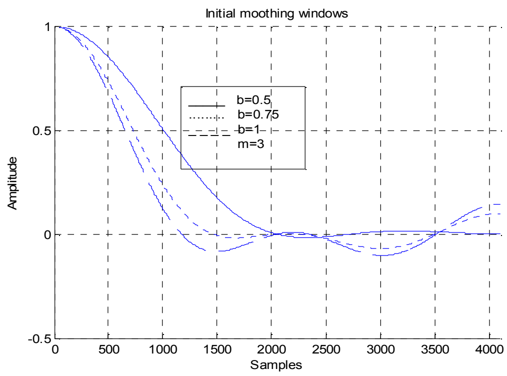
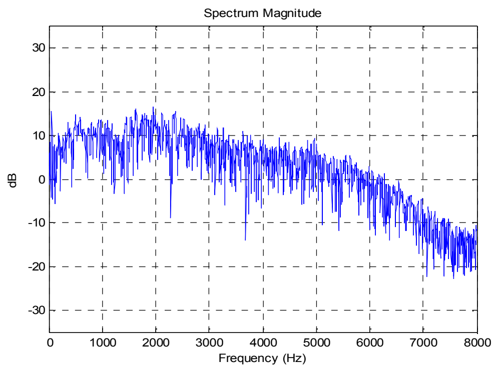



Fig 2 Initial smoothing windows for $i = 0$

(a)




(b)

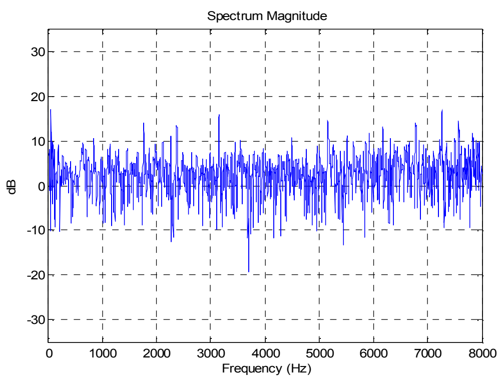



Fig 5 -1 Deviation of magnitude response from flatness
(a) Initial original response, b) Equalized response

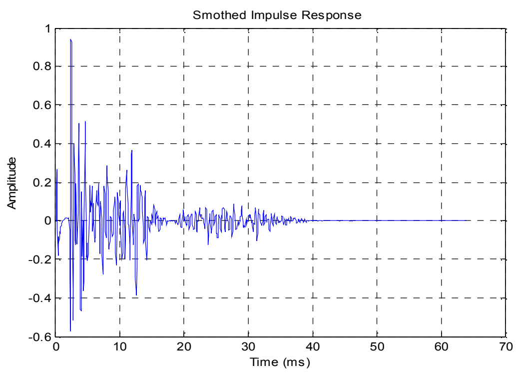



Fig 3 Smoothed impulse response $h_{cs}^{3}(n)$

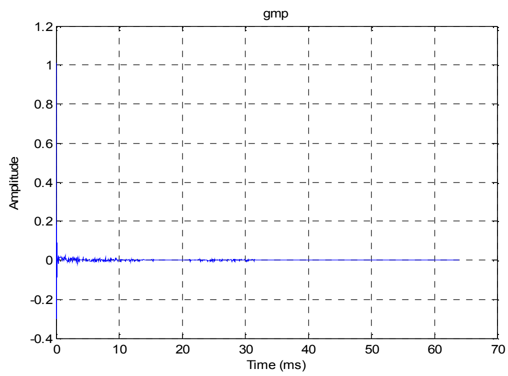
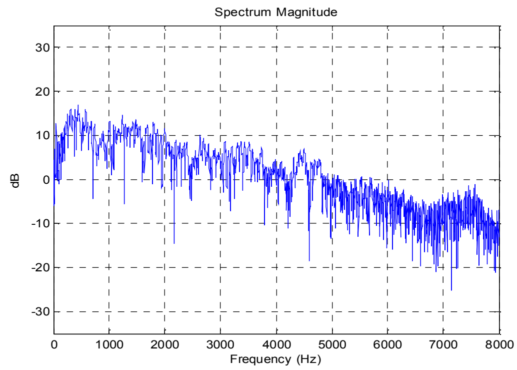



Fig 4 Partial minimum-phase inverse impulse response $g_{cs,mp}^{1/8}(n)$ calculated using the iterative version of the homomorphic method with $L = 8$ .

(a)




(b)

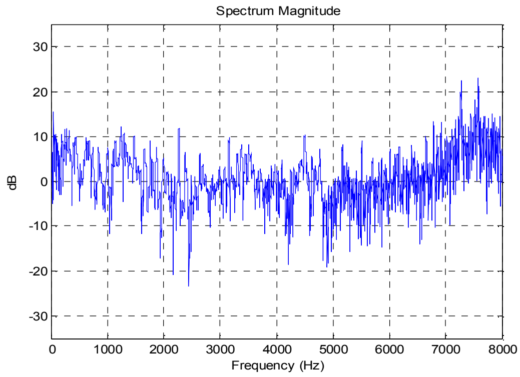



Fig 5-2 Deviation of magnitude response from flatness
a) Second original response, b) Equalized response

## V. IMPLEMENTATION OF AN EQUALIZATION FILTER ON TMS320VC5510 DSK DSP PLATFORM

An equalization filter representing the example of Section 4 (Figure 7) with reduced complexity (significantly about 1000 samples rather than more than 10000 required in a complete equalization – perfect dereverberation) has been implemented on TMS320VC5510 DSK kit (Figure 8). We used a version 2 of Code Composer Studio to load on the DSP platform the C source coefficients (coded 16 bits integer – int16) of the filter generated by MATLAB and execute the C source program of an FIR filter implementation. The test signals (wave files coded 16 bits at 16 kHz) transmitted from PC used in this application for listening evaluations were the same used in Section 4, but passed trough AIC23 Codec of the DSP kit (for processing from the line input (LINE IN) to the line output (LINE OUT)). If DIP switch 3 is high no filtering of audio Data, but if DIP switch 3 is down audio Data are filtered by the FIR coefficients before playing the results from LINE OUT of AIC23 Codec.

The first signal is a reverberated speech, obtained after convolution of an anechoic speech with an initial acoustic impulse response of an audio-conferencing room (Fig 1 and corresponding to magnitude response of Figure 5-1- (a). The second signal is also a reverberated speech, but obtained after convolution of the same anechoic speech with a second acoustic impulse response which is different than the initial one (corresponding to magnitude response of Figure 5-2 – (a), representing in this way possible changes in the listener position inside the room. The third and fourth signal are respectively the equalized (filtered) first and second signal (first and second signal passed trough DSP module for filtering by FIR coefficients - when DIP switch 3 is down - before playing the results from the output (LINE OUT) of AIC23 Codec.

These four test signals recorded by MATLAB's Audio Tools from LINE OUT of AIC23 Codec were presented binaurally in randomly order to young listeners (about 25 graduation students in a quiet listening class room) with normal hearing through a headphone (Logitech – Premium Stereo Headset) at a comfortable level. The listening tests evaluation indicates in Figure 9 as it was expected in Section 4, a significant perceptual preference for the equalized signals over the original signals (reverberated speech) for both cases, the first and second signal (that is, respectively the third and fourth signal). These results can be confirmed in an anechoic chamber by transmitting the recorded wave files trough an adapted loudspeaker to some listeners placed inside.

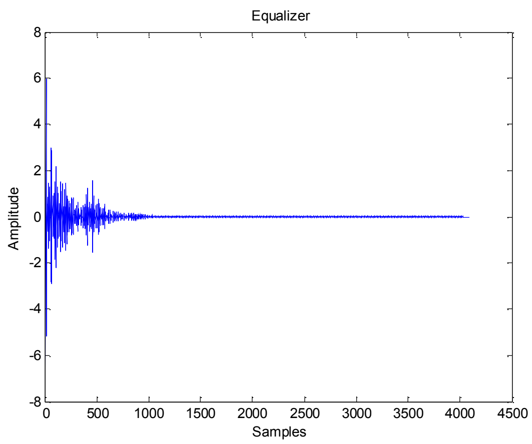



Fig 7: Impulse response of the equalization filter obtained after 8 successive convolutions of partial minimum-phase inverse impulse response of the example shown in the Figure 4, section 4.

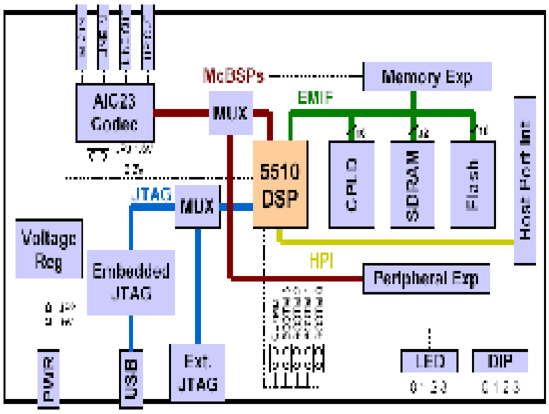
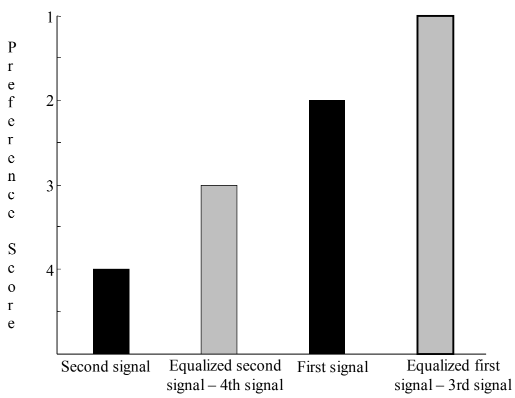



Fig 8: Block diagram VC5510 DSK

*Key Features*:
- A Texas Instruments 5510 DSP operating at 200MHz
- An AIC23 stereo codec
- 8 Mbytes of synchronous DRAM
- 512 Kbytes of non-volatile Flash memory
- 4 user accessible LEDs and DIP switches
- Software board configuration through registers implemented in CPLD
- Jumper selectable boot options
- Standard expansion connectors for daughter card use
- JTAG emulation through on-board JTAG emulator with USB host interface or external emulator
- Single voltage power supply (+5V)




Preference Score 1: the most preferred signal.

Fig 9: Listening tests evaluation

## VI. CONCLUSIONS

A technique for dereverberation systems design based on partial equalization for room acoustics is presented. In this technique, a modification of the corresponding long duration measured acoustic impulse response is produced by applying an iterative simple smoothing process. Corresponding impulse responses of reduced complexity are then produced, but preserve the initial transient data that are significant to the listener (direct signal and first reflections). The algorithm also gives the possibility of stopping the smoothing process when a corresponding designed equalization filter (inverse filter) produces some pre-determined equalized magnitude response. Equalization results have been validated in the context of an audio-conferencing room for a reduced complexity equalization filter (FIR inverse filter) using listening test with a headphone, MATLAB audio tools and implementation on Texas Instruments DSP board

## REFERENCES


[1] A.Gilloire et J.P.Jullien, 'L`acoustique des salles dans les télécommunications', L`écho des RECHERCHES, No 127, 1er trimestre 1987.

[2] Neely S. T. and J. B. Allen, "Invertibility of a room impulse response", J. Acoust. Soc. Am, vol. 66, no. 1, pp. 165-169, July 1979.

[3] B. D. Radlovic and R. A. Kennedy, "Non-minimum phase equalization and its subjective importance in room acoustics", IEEE Transactions on Speech and Audio Processing, vol. 8, no. 6, Nov 2000.

[4] J. N. Mourjopoulos, "Digital Equalization of Room Acoustics", J. Audio. Eng. Soc, Vol 42, N°11, 1994 November.

[5] Maamar Ahfir, Izzet Kale, Berkani Daoud, "Room Equalization Based on Iterative Simple Complex Smoothing of Acoustic Impulse Responses" IEEE ICASSP 2006, Toulouse (FRANCE), Vol V, pp 109-112, 17-21 May 2006.

[6] P.Hatziantonio and J.Mourjoupoulos, " Generalized Fractional- Octave Smoothing of Audio and Acoustic Responses", J. Audio. Eng. Soc, vol 48, No 4, pp 259-280, 2000 April.

[11] J..P.Jullien, A.Gilloire and A.Saliou, "Mesure des réponses impulsionnelles en acoustique", Note technique NT/LAA/TSS/181, Novembre 1984.

[12] B. D. Radlovic and R. A. Kennedy, "Iterative cepstrum-based approach for speech dereverberation ", ISSPA` 99, Brisbane, Australia, 22-25 August 1999.

[7] J.N.Mourjopoulos, "Digital Equalization of Room Acoustics", J. Audio. Eng. Soc, Vol 42, N°11, 1994 November.

[8] P.Hatziantonio and J.Mourjoupoulos, " Generalized Fractional-Octave Smoothing of Audio and Acoustic Responses", J. Audio. Eng. Soc, vol 48, No 4, pp 259-280, 2000 April.

[9] P. M .Clarkson, J.Mourjopoulos and J.K.Hammond, "Spectral, Phase and Transient Equalization for Audio Systems", J. Audio. Eng. Soc, vol 33, pp 127-132, 1985 Mar.

[10] S. Salamouris, K. Politopoulos, V. Tsakiris and J. Mourjopoulos, "Digital System for Loudspeaker and Room Equalization ", presented at the 98th Convention of the Audio Engineering Society, J. Audio. Eng. Soc, (Abstracts), vol 43, pp 396, 1985 May.