

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Departamento de Engenharia Electrotécnica e de Computadores

IMPLEMENTAÇÃO EM DSP DE UM SISTEMA “REAL-TIME ANALYZER”

APLICAÇÃO À IGUALIZAÇÃO ADAPTATIVA

Gabriel Falcão Paiva Fernandes

**Licenciado em Engenharia Electrotécnica e de Computadores
pela
Faculdade de Engenharia da Universidade do Porto**

**Dissertação submetida para satisfação parcial dos
requisitos do grau de mestre**

**em
Engenharia Electrotécnica e de Computadores
(Área de especialização em Telecomunicações)**

**Dissertação realizada sob a supervisão do
Professor Doutor Aníbal João de Sousa Ferreira,
do Departamento de Engenharia Electrotécnica e de Computadores
da Faculdade de Engenharia da Universidade do Porto**

Porto, Março de 2002



RESUMO

Todas as salas produzem uma resposta acústica quando excitadas por uma fonte sonora. O objectivo principal deste trabalho é encontrar um meio computacional baseado em plataforma DSP, de caracterizar essa mesma resposta acústica, compensando em tempo-real as anomalias introduzidas pela sala na reprodução do som.

Em particular, este trabalho explora formas alternativas de extrair essa resposta acústica, apresenta uma técnica prática e eficiente de realizar filtragem FIR rápida no domínio das frequências, e detalha o sistema desenvolvido para implementar igualização adaptativa.

Este sistema é baseado numa plataforma de processamento – *SHARC ADSP-21065L EZ-LAB*, e numa outra de aquisição de sinais áudio de elevada qualidade – *AD1836 Evaluation Board*, ambas da *Analog Devices*. Pretende-se que o sistema desenvolvido seja capaz de, em tempo-real, excitar uma sala com uma determinada onda sonora, capturar a resposta daquela a esse mesmo sinal áudio através de um microfone e recalculer e actualizar os novos parâmetros de igualização de forma adaptativa, tendo por objectivo uma determinada resposta em frequência desejada para um ponto da sala centrado no microfone.

ABSTRACT

Every single room produces a certain acoustical response when excited by a sound source. The main purpose of this work is to develop a computational solution based on a DSP platform, to the problem of characterizing that acoustical response in order to perform a real-time compensation of anomalies introduced by the room in sound reproduction.

In particular, this work explores alternative ways of extracting that acoustical response, presents a practical and efficient technique of implementing FIR fast filtering in the frequency domain, and details the system developed to produce adaptive equalization.

This system is based on a signal processing platform – *SHARC ADSP-21065L EZ-LAB* and on a high quality audio acquisition board – *AD1836 Evaluation Board*, both from *Analog Devices*. The developed system should be able to perform, in real-time, the excitation of the room with a sound wave, the capture of the response to that same signal using a microphone, the adaptive calculation and update of the new equalization settings in order to obtain a desired frequency response for a region in the room centered at the microphone.

PRÉCIS

Toutes les salles produisent une réponse acoustique lorsqu'elles sont excitées par une source sonore. L'objectif principal de ce travail c'est de rencontrer un moyen computationnel pour caractériser cette réponse acoustique, basé en une plate-forme DSP, de façon à compenser en temps-réel, les anomalies introduites par la salle dans la reproduction du son.

Spécifiquement, ce travail explore des façons alternatives d'obtenir cette réponse acoustique, il présente une technique pratique et efficiente pour accomplir un filtrage FIR rapide dans le domaine des fréquences, et il détaille le système développé pour implémenter l'égalisation adaptative.

Ce système-ci est fondé dans une plate-forme de traitement - *SHARC ADSP-21065L EZ-LAB*, et dans une autre d'acquisition des signaux d'audio de haute qualité - *AD1836 Evaluation Board*, tous les deux provenant de l'*Analog Devices*. On veut démontrer que le système développé est capable d'exciter une salle en temps-réel avec une onde sonore déterminée, qu'il est capable de capter la réponse de la salle à ce même signal audio à travers d'un microphone, et encore qu'il peut recommencer à calculer et à actualiser les nouveaux paramètres d'égalisation d'une façon adaptative, ayant pour objectif une réponse déterminée en fréquence souhaitée pour une région de la salle centré au microphone.

PREFÁCIO

As actuações ao vivo são parte integrante de todas as sociedades humanas e os problemas associados à distribuição do som a partir de um pequeno número de falantes para um elevado número de ouvintes, são conhecidos desde há alguns séculos. Nos teatros da Grécia antiga, a audiência sentava-se em lugares de inclinação íngreme, em torno de um palco onde a voz dos actores radiava com suficiente amplitude sobre toda a plateia. Eram colocadas superfícies reflectoras de som em locais próximos do palco (tipicamente chão de pedra), que permitiam uma melhor propagação da voz dos actores sobre a audiência. Estes usavam máscaras, que possuíam embocaduras capazes de amplificar a voz e muitas vezes declamavam a partir de caixas ressonantes.

Posteriormente, nos grandes espaços fechados da Idade Média, como eram as igrejas, a reverberação provocada pelas paredes reflectivas de pedra, era responsável por distribuir o som facilmente a todos os elementos da audiência e proporcionar um reforço musical. Por outro lado, tempos de reverberação muito longos tornavam a fala difícil de compreender. Como resultado, os cânticos tinham de ser efectuados de forma lenta, para se tornarem perceptíveis, daí, pensa-se, o aparecimento do canto Gregoriano e outros arranjos corais executados de forma lenta. A concepção de teatros, que se tornaram abundantes no período da Renascença, era baseada em princípios Aristotélicos e não científicos, pelo que aqueles que produziam boa acústica, tornavam-se modelos de referência para futuros teatros.

A era da acústica moderna começa no princípio do século XX com *W. C. Sabine* e as suas experiências que levaram à publicação da famosa fórmula de reverberação, que relaciona o volume de uma sala com a sua área de absorção. A partir deste trabalho, foi-se adquirindo a noção de que há outros parâmetros relevantes além de reverberação, que contribuem igualmente para a impressão subjectiva da qualidade de som numa sala. Muitos investigadores das diferentes áreas da engenharia (e.g. Electrotecnia, Civil) têm dedicado esforços ao seu estudo, reflectidos em inúmeras publicações na área, com o objectivo de melhor compreender os princípios pelos quais se regem, o que lhes tem permitido projectar sistemas acústicos de raiz, cada vez mais perfeitos, ou melhorar aqueles que pelos mais diversos motivos não o são.

Com o advento das novas tecnologias de processamento digital de sinal, durante a última década, surgiram meios computacionais que permitem manipular sinais áudio digitais de forma extremamente eficiente a custos reduzidos. Fazendo uso dessas tecnologias, a presente dissertação propôs-se encontrar um meio computacional rápido e eficiente, capaz de realizar o estudo de determinados parâmetros acústicos e de encontrar formas adaptativas e automáticas de os melhorar, em tempo-real.

Concretamente, pretende-se caracterizar a acústica de uma sala, influenciada entre outros fenómenos pelas reflexões do som nas suas superfícies, e posteriormente corrigi-la, aproximando-a a uma resposta em frequência desejada, devendo estas tarefas ser realizadas num curto espaço de tempo.

AGRADECIMENTOS

Foi no *Grupo de Áudio da Unidade de Telecomunicações do INESC–Porto* que cresceu o gosto pelo processamento de sinal e foi nele que tive a oportunidade de desenvolver esta tese.

Um trabalho deste género só foi possível graças ao apoio de muitas pessoas, que das mais diversas formas me incentivaram ao longo da sua realização.

Ao meu orientador, o Prof. Aníbal Ferreira quero agradecer o apoio, os inúmeros conselhos e a revisão exaustiva deste texto, bem como o entusiasmo que sempre demonstrou por este trabalho. A ele devo o facto de poder fazer investigação nesta área. Ao Miguel e ao Gustavo, deixo a minha gratidão pela amizade, pelas opiniões e sugestões ao meu trabalho. Com eles e ao longo de mais de dez anos de trabalho em conjunto, adquiri o conhecimento e a experiência que me permitem concluir hoje este trabalho. Uma palavra de apreço pela forma paciente com que conviveram com os ensaios ruidosos, tão frequentes nesta sala durante os últimos meses.

Gostava de agradecer ao Prof. Oliveira de Carvalho a pronta disponibilidade na partilha de conhecimentos e o entusiasmo com que sempre encarou as nossas conversas sobre acústica. Muitos dos problemas relacionados com esta área e que foram surgindo no decorrer do trabalho, ultrapassaram-se graças às suas valiosas sugestões.

Queria igualmente demonstrar a minha gratidão a todo o *Grupo de Opto-Electrónica do INESC–Porto*, em particular aos Eng.^{os} Filipe Pinto e Alberto Maia, pelos trabalhos desenvolvidos em conjunto e que tanto me têm dado a aprender.

Uma palavra de apreço ao Prof. Sílvio Abrantes pelo apoio em processamento adaptativo.

Aos meus amigos, cuja presença é por si só factor motivador, quero agradecer o incentivo, que me permitiu levar esta tarefa a bom porto.

Ao meu irmão e melhor amigo, agradeço a amizade e crítica sinceras, que em tantos momentos trazem a lucidez necessária.

Este trabalho nunca teria sido possível sem o apoio da minha mãe e do meu pai, principalmente nos momentos difíceis. A confiança inabalável com que sempre encararam esta tarefa, bem como tantas outras, é o combustível dos sucessos que temos alcançado. A eles, um obrigado que não sei expressar. A eles, dedico esta tese.

Agradecendo a amizade e dedicação sinceras, uma parte deste trabalho é também dedicada à minha outra mãe, Ana Maria.

Esta dissertação tem sido uma demonstração da tua paciência, Gui. Agradecer-te não me parece suficiente. Aliás, nada mo parece. Lembra-te apenas: tu deste-me o motivo...

NOTAS PARA O LEITOR

A fonte usada em texto normal é Times New Roman. Para código fonte usa-se o tipo Courier New. No início estão listados os acrónimos usados ao longo do texto, de forma a não o sobrecarregar com excessivas definições. As referências bibliográficas estão listadas no final do documento, sendo indicadas pelo último nome do autor seguido do ano de publicação dentro de parêntesis rectos.

ÍNDICE

Resumo.....	1
Abstract	2
Précis.....	3
Prefácio	4
Agradecimentos	5
Notas para o Leitor	6
Índice.....	7
Lista de Figuras	9
Lista de Tabelas	10
Acrónimos	11
1. Introdução.....	13
1.1. Introdução Histórica	13
1.2. Sistemas existentes	14
1.3. Aplicações.....	15
2. Caracterização do problema	17
2.1. Identificação do problema	17
2.2. Objectivos	20
3. Critérios em acústica de salas	21
3.1. Medições.....	21
3.2. Modos de Ressonância da Sala	23
3.3. Loudness	24
3.4. Tempo de Reverberação	25
3.5. EDT – Early Decay Time	26
3.6. Claridade.....	26
3.7. Definição	26
3.8. Tempo central (Center Time)	27
3.9. Força.....	27
3.10. LEF – Lateral Energy Fraction	28
4. Correção da Resposta Acústica da Sala.....	29
4.1. Tipos de Excitação da Sala	29
4.1.1. Impulsiva.....	29
4.1.2. Aleatória.....	30
4.1.3. Sinais Musicais.....	30
4.1.4. Varrimento Discreto de Sinusóides (stepped sine).....	30
4.1.5. Sequências de Comprimento Máximo (MLS)	31
4.2. Filtragem Adaptativa.....	32
4.2.1. Extracção de Características	32
4.2.2. Método LMS standard	33
4.2.3. Analogia entre algoritmo implementado e LMS <i>standard</i>	40
4.3. Cálculo Adaptativo de Novos Parâmetros de Igualização	44
4.3.1. Excitação impulsional por disparo de pistola sonora	44
4.3.2. Algoritmo com excitação por ruído branco	45
4.3.3. Algoritmo com excitação através de sequências MLS	58
5. Sistema Igualizador.....	75
5.1. Diagrama de Blocos do Sistema	75
5.2. Parâmetros de Igualização	79

6. Implementação em DSP em Tempo-Real	80
6.1. Hardware	80
6.1.1. ADSP SHARC 21065L Analog Devices (processamento em tempo-real)	81
6.1.2. Placa de aquisição áudio AD1836.....	84
6.1.3. Sistema de Amplificação, Microfone e Colunas	85
6.2. Software.....	86
6.2.1. Implementação mixed C/Assembly.....	87
6.2.2. Optimização de código para operação em tempo-real.....	90
6.2.3. Rotina de Atendimento a Interrupções	92
7. Testes e Resultados	93
7.1. Realização de Testes em Sala com condições Acústicas Desfavoráveis	93
7.2. Apreciação Subjectiva	93
8. Conclusões.....	94
Bibliografia.....	96

LISTA DE FIGURAS

Figura 2.1 - Planta da sala	17
Figura 2.2 - Reflexões nas superfícies da sala. a) Representação de 170ms do sinal, b) Pormenor dos primeiros 17ms para visualização das primeiras reflexões.	18
Figura 2.3 – Som directo e primeiras reflexões	19
Figura 3.1 - Reverberação	25
Figura 4.1 - Sistema Linear e Invariante nos Tempos.....	33
Figura 4.2 - Filtragem adaptativa	34
Figura 4.3 - Performance da superfície de erro para filtro de dois coeficientes h_0 e h_1	38
Figura 4.4 - Modelo da filtragem adaptativa usada na realização do trabalho	41
Figura 4.5 – Resposta da sala a um disparo de pistola	44
Figura 4.6 - Esquema de excitação por ruído branco	46
Figura 4.7 - Ajuste adaptativo de novos parâmetros de igualização	47
Figura 4.8 – Ajuste em torno de MAX e MIN de $y(k)$	49
Figura 4.9 - Ajuste em torno da média aritmética de $y(k)$	51
Figura 4.10 - Resposta em frequência inicial da sala.....	52
Figura 4.11 - Curva de aprendizagem para o método 4.2.4.2.....	53
Figura 4.12 – Níveis de igualização da sala obtidos adaptativamente	54
Figura 4.13 - Resposta em frequência final da sala.....	56
Figura 4.14 - Mapeamento da resposta em frequência nas 20 bandas	57
Figura 4.15 - <i>Shift register</i> gerador de sequências MLS	59
Figura 4.16 - Conversão de uma sequência MLS para o domínio contínuo.....	64
Figura 4.17 - Algoritmo MLS em linguagem C.....	65
Figura 4.18 - Ajuste adaptativo de novos parâmetros de igualização com excitação MLS	66
Figura 4.19 - Resposta impulsional da sala para os primeiros 40ms	68
Figura 4.20 - Curva de aprendizagem para o método 4.2.4.3.....	69
Figura 4.21 - Resposta inicial da sala para o método 4.2.4.3	70
Figura 4.22 - Níveis de igualização da sala obtidos adaptativamente para o método 4.2.4.3.....	71
Figura 4.23 - resposta final da sala para o método 4.2.4.3	72
Figura 4.24 - Resposta em frequência nas 20 bandas para o método 4.2.4.3	73
Figura 5.1 – Diagrama de blocos do sistema	76
Figura 5.2 – Aproximação aos parâmetros de igualização desejados	79
Figura 6.1 – Diagrama de blocos do ADSP-21065L Analog Devices	82
Figura 6.2 – Formato de trama do ADC.....	85
Figura 6.3 – Resposta em frequência do microfone TR40 Audix.....	86
Figura 6.4 – Butterfly da FFT otimizada para SHARC 21065L.....	91

LISTA DE TABELAS

Tabela 1 - Níveis de igualização final	55
Tabela 2 - Níveis de igualização para método MLS	71

ACRÓNIMOS

DUT	Device Under Test
DSP	Digital Signal Processor
IR	Impulse Response
ODFT	Odd-discrete Fourier Transform
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
FHT	Fast Hadamard Transform
ISR	Interrupt Service Routine
DAG	Data Address Generator
DMA	Direct Memory Access
EDT	Early Decay Time
LEF	Lateral Energy Fraction
MFLOP	Million Floating Point Operation
SHARC	Super Harvard Architecture
PM	Program Memory
DM	DataMemory
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
LMS	Least Mean Square
FDAF	Frequency-Domain Adaptive Filtering
FIR	Finite Impulse Response
MLS	Maximum Length Sequence
ADC	Analog-to-Digital Converter
DAC	Digital-to-Analog Converter

"There are many hypotheses in science which are wrong. That's perfectly all right; they're the aperture to finding out what's right. Science is a self-correcting process. To be accepted, new ideas must survive the most rigorous standards of evidence and scrutiny."

Carl Sagan, COSMOS

1. INTRODUÇÃO

As descobertas realizadas na área de Processamento de Sinal durante as últimas décadas, conduziram ao desenvolvimento de novas tecnologias baseadas em processadores rápidos, com grandes quantidades de memória (de acesso rápido) disponível, possuidores de inúmeros periféricos especialmente dedicados à implementação e optimização de algoritmos para processamento de sinal – nomeadamente de sinal áudio e vídeo –, como sejam canais de DMA, portas série, memória *Dual Port*, etc. Acompanhando este processo surgiram ferramentas de desenvolvimento novas e poderosas direccionadas ao mercado de DSPs, com bom suporte ao nível de linguagem de programação, nomeadamente:

- no fornecimento de vastas bibliotecas de funções optimizadas, especialmente desenvolvidas para aplicações de processamento de sinal;
- na facilidade de utilização de código *Assembly* juntamente com C;
- e nos ambientes de *debugging* amigáveis e repletos de funcionalidades úteis.

Se a todas estas vantagens, associarmos o facto de que estes sistemas passaram a estar nos últimos anos disponíveis no mercado a custos reduzidos, verifica-se que estão mais do que nunca criadas as condições para implementar em DSP – fazendo portanto uso de tecnologia digital com todas as vantagens que daí advêm –, muitas soluções que dificilmente o seriam recorrendo a tecnologia analógica, ou que quando implementadas nesta tecnologia apresentam desvantagens consideráveis (e.g. igualização).

Como consequência, é hoje possível desenvolver e experimentar com rapidez novas soluções baseadas em algoritmos mais elaborados, fortemente direccionados para as crescentes necessidades na área da implementação em tempo-real de algoritmos de processamento de sinal. Estas experiências eram dificilmente realizáveis há alguns anos, se não de todo impossíveis.

Graças à situação actual, foi-nos possível desenvolver a “Implementação em DSP de um sistema Real-Time Analyzer”, que tem como principal função a análise, caracterização e correcção das condições de reprodução sonora num espaço acústico.

1.1. INTRODUÇÃO HISTÓRICA

A acústica é a ciência que estuda a produção, transmissão e efeitos do som. O som não é apenas um fenómeno que provoca a sensação de audição através do ar. Por esse motivo, frequências inferiores (infrasons) ou superiores (ultrasons) à gama audível por um ser humano são também entendidas como som.

O som é uma forma de transmissão de informação. A transmissão de som é afectada por, e consequentemente dá informação sobre, os meios que percorre e os obstáculos que encontra, como por exemplo, as heterogeneidades de uma sala ou auditório.

A especulação de que o som é um fenómeno ondulatório surgiu com a observação da propagação de ondas na água. Uma definição primária de onda era dada como sendo uma perturbação oscilatória que se afasta de uma determinada fonte, transportando uma quantidade de matéria não discernível por grandes distâncias de propagação. A possibilidade do som exibir um comportamento análogo foi discutida por exemplo pelo filósofo grego *Chrysippus* (240 A.C.), pelo arquitecto e engenheiro romano *Vetruvius* (25 A.C.) e pelo filósofo romano *Boethius* (480-524) [Pierce81]. Essa interpretação era também consistente com a de *Aristóteles* (384-322 A.C.), quando este afirmava que o movimento do ar era gerado por uma fonte que “empurrava para a frente o ar adjacente, de maneira a este viajar com a sua qualidade inalterada enquanto as perturbações no ar o permitissem”. Uma experiência bastante conclusiva relatada em *Harmonic Universelle* (1636) por *Marin Mersenne* (1588-1648), muitas vezes referido como o “pai da acústica”, demonstrou que a relação de frequências entre um tom puro gerado por uma corda e outro tom uma oitava acima, é de 1:2. Ora, a percepção dos dois tons ficaria explicada se a relação entre as duas ondas oscilantes no ar fosse também de 1:2. Esta afirmação é também consistente com a supra referida teoria da *fonte-ar-movimento-frequência*. *Galileo Galilei* (1564-1642) assume também bastante protagonismo nesta área uma vez que o seu “*Mathematical Discourses Concerning Two New Sciences*” (1638) contém o discurso mais lúcido dado até à época sobre a equivalência de frequência.

A teoria matemática de propagação do som começou com *Isaac Newton* (1642-1727) cujo “*Principia*” (1686) incluía uma definição de som como sendo variações pulsatórias na pressão transmitidas através de partículas vizinhas num fluido.

Progressos substanciais na teoria de propagação do som foram feitos no século XVIII por *Euler* (1707-1783), *Lagrange* (1736-1813) e *D’Alembert* (1717-1783), baseados em conceitos matemáticos e físicos mais sólidos. A teoria proposta para o som encontrava-se incompleta em muitos pontos de vista, se bem que a maioria das teorias modernas possa ser vista como um refinamento das desenvolvidas por *Euler* e pelos seus contemporâneos.

Finalmente, com o aparecimento de *W. C. Sabine* e da sua fórmula que relaciona a reverberação de uma sala com o seu volume e área de absorção, nasceu em muitos investigadores a sensibilidade para a necessidade de aprofundamento do estudo sobre este parâmetro, bem como de vários outros que foram sendo descobertos, e que contribuem para a impressão subjectiva da qualidade do som de uma sala.

1.2. SISTEMAS EXISTENTES

Com base numa pesquisa efectuada, concluímos que há diversos equipamentos comercialmente disponíveis, capazes de efectuar uma extracção de características acústicas de uma sala e de efectuar determinados tipos de correcções, dos quais se ilustram alguns exemplos.

Peavey CEQ 280a

O equipamento *CEQ 280a* fabricado pela empresa *Peavey* possui um *Real-Time Analyzer* (RTA) incorporado e uma fonte de ruído rosa para auto-igualização de uma sala (<http://www.peavey.com>). Usando esta fonte sonora para gerar o sinal excitador, é possível calcular a resposta em frequência do sistema. A resposta em frequência pode ser obtida através de sucessivas capturas em diferentes pontos da sala, sendo que a final será uma média das anteriores, pelo que a curva de igualização obtida deverá ser usada apenas para uma zona restrita da sala (por exemplo, o sofá onde se encontra o ouvinte), já que a resposta em frequência pode variar bastante para diferentes pontos de uma mesma sala.

Behringer Ultra-curve Pro DSP8024

O sistema RTA do equipamento *Ultra-curve Pro DSP8024* permite medidas de $\frac{1}{3}$ de oitava com precisão de 0.25dB (<http://www.behringer.com>). Permite a geração de sinais excitadores do tipo ruído branco, ruído rosa e sinusóides ajustáveis em passos de 1/60 oitava. Permite, de igual modo, medir a resposta do sistema e proceder de seguida à actualização automática dos parâmetros de igualização. Possui ainda funcionalidades importantes que impedem a atribuição de parâmetros de igualização impraticáveis. Se por exemplo, o sistema emitir um aviso a um pedido de ganho superior a um determinado valor, para uma determinada banda atingir a curva desejada, isso pode significar que a coluna não é capaz de reproduzir essa frequência (tipicamente nas muito baixas ou altas frequências). Neste caso, o sistema não efectua qualquer alteração no ganho da banda evitando a sobrecarga das colunas.

SigTech AEC-1000 Room Acoustics Equalizer

O equipamento *AEC-1000 (Acoustic Environmet Correction)* é constituído por um igualizador digital, direccionado para o mercado profissional (estúdios, auditórios, etc.), embora também possua uma versão para consumo doméstico. Permite armazenar até quatro modos de correcção (para diferentes colunas ou diferentes locais de audição na sala) em tempo-real, e como os dois sistemas descritos anteriormente, necessita de um microfone para captar a resposta da sala. Este necessita, no entanto, adicionalmente de um PC para correr o software de síntese do filtro e o descarregar para o AEC-1000. Durante a fase de calibração da sala (cálculo do filtro), o sistema excita-a com impulsos, que são alternadamente emitidos por cada uma das duas colunas de som. Comparando a resposta da sala a esses impulsos com os impulsos originais, é possível obter a sua resposta e a partir daí, actuar no sentido de corrigir ecos e reflexões que produzam efeitos indesejados no som.

1.3. APLICAÇÕES

O principal objectivo do trabalho descrito nesta dissertação é desenvolver um sistema capaz de efectuar correcções em áudio, eliminando perturbações introduzidas nas condições de reprodução para um determinado local da sala, devidas por exemplo a reflexões do mesmo nas

superfícies da sala em análise. A sua utilização é indicada em locais onde um determinado orador se pretenda fazer ouvir por uma pessoa, ou por um conjunto de pessoas, como é o caso de igrejas, escolas, auditórios, centros cívicos e outras instalações de tamanho pequeno ou médio. Aplica-se de igual modo em situações onde, para os espaços referidos, se pretenda fazer uso de sistemas de reprodução musical, como por exemplo na sala de uma casa. Tipicamente, procura-se aplanar a resposta em frequência da sala em estudo para um determinado local de audição (onde será colocado o microfone), a fim de eliminar deformações provocadas no espectro por reflexões.

2. CARACTERIZAÇÃO DO PROBLEMA

O trabalho realizado na implementação, teste e validação dos sistemas descritos neste texto, foi realizado no Laboratório de Processamento de Sinal da Faculdade de Engenharia da Universidade do Porto. As características da sala (ou DUT – *Device Under Test*), bem como os efeitos que introduz na qualidade do som que nela é reproduzido, foram estudados e são descritos de seguida. É dada grande ênfase ao estudo das reflexões do sinal acústico nas superfícies da sala, que são em grande parte, como se demonstra neste capítulo, responsáveis pela alteração da qualidade do som.

2.1. IDENTIFICAÇÃO DO PROBLEMA

A sala na qual se efectuaram todos os testes, possui 10m de comprimento por 6m de largura e 3m de altura. É um laboratório ocupado por várias secretárias, estantes e armários, dispostos da forma indicada na Figura 2.1. Nele trabalham permanentemente três pessoas, o que obrigou a que todos os testes se realizassem durante o período nocturno, a fim de evitar interferências nas medidas.

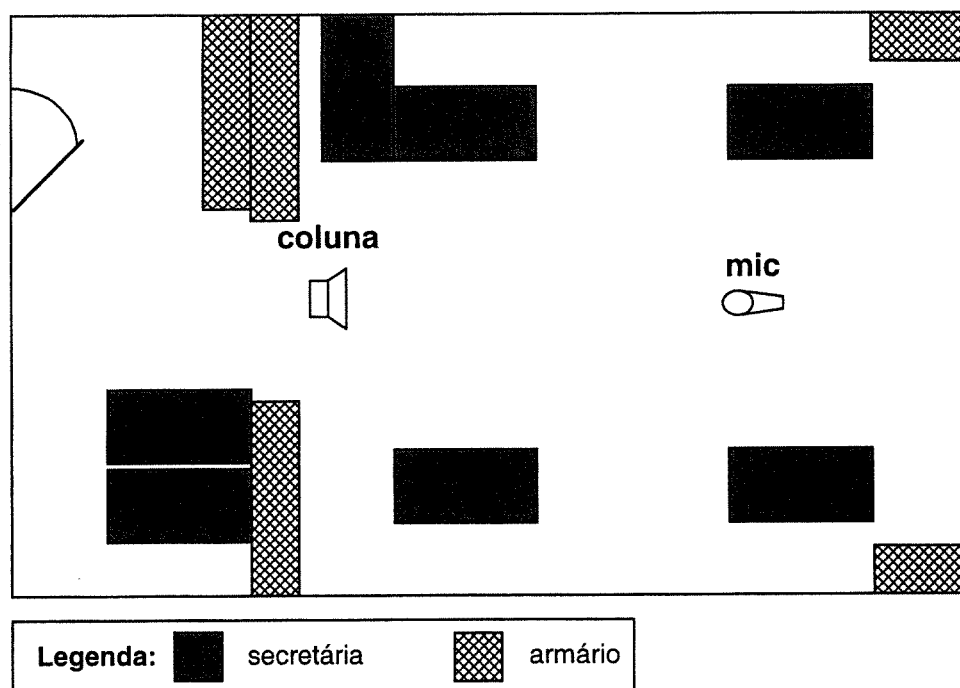


Figura 2.1 - Planta da sala

Não há obstáculos entre a coluna de som e o microfone. Desta forma, simula-se uma sala de audição, onde a zona do microfone equivale àquela que seria a localização do ouvinte. Quer a coluna de som, quer o microfone, situam-se 1.26m acima do nível do solo. Por este motivo, a primeira reflexão do sinal usa o chão, enquanto que a segunda se reflecte no tecto (a 1.74m) e as seguintes nas superfícies laterais (a 3m). Estas últimas revelam-se razoavelmente reflectoras, uma vez que lhes foi aplicado um género de tinta plástica.

Reflexões nas superfícies da sala

Testes efectuados à sala a partir da excitação desta com um sinal semelhante a um impulso (diferente na realidade, já que era necessário utilizar um sinal com energia suficiente para excitar os modos da sala), permitiram medir o atraso entre o sinal directo e as primeiras reflexões. A Figura 2.2 a) representa a resposta da sala ao impulso durante 170ms, dando uma indicação da cauda de decaimento de energia. A Figura 2.2 b) mostra a anterior com detalhe nos primeiros 17ms, e nela são claramente visíveis as primeiras reflexões do sinal provocadas pelas superfícies da sala. Essas reflexões são responsáveis pelos reforços e cancelamentos introduzidos no sinal às mais diversas frequências, para o ponto de audição (correspondente ao microfone). Elas produzem a resposta da sala para aquele ponto. Como referido anteriormente, verifica-se que a primeira reflexão é proveniente do chão. O tempo de atraso directo¹ medido (que não é mostrado na Figura 2.2), indicou 570 amostras. Sabendo a frequência de amostragem – 48kHz, equivale a dizer que no seu percurso directo, o som percorre 4.07m – *dist1*, em 11.86ms (à velocidade do som – que a 20°C é cerca de 343ms⁻¹).

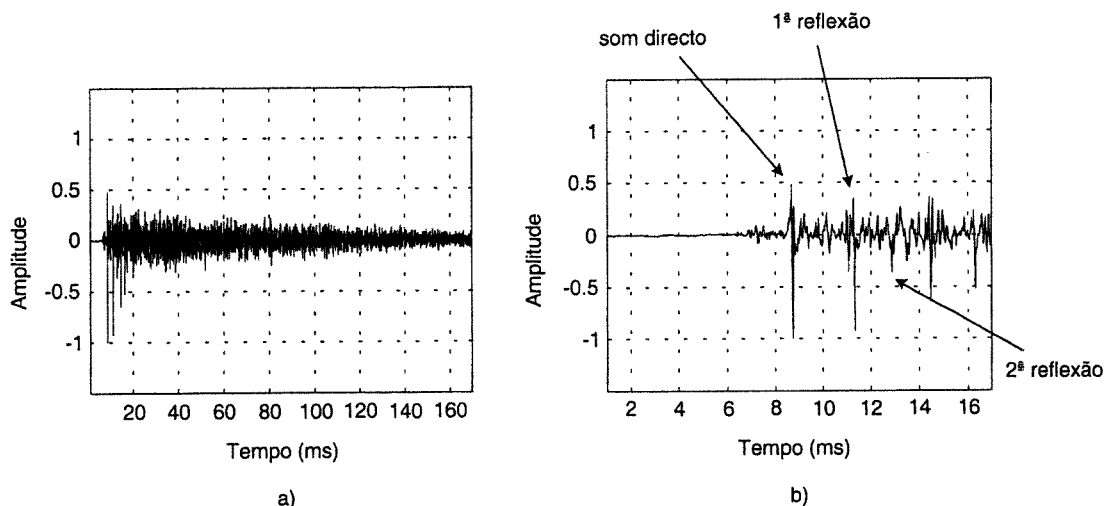


Figura 2.2 - Reflexões nas superfícies da sala. a) Representação de 170ms do sinal, b) Pormenor dos primeiros 17ms para visualização das primeiras reflexões.

Observe-se agora o cálculo da primeira reflexão, com a ajuda da Figura 2.3.

¹ Tempo mínimo que o sinal demora a percorrer a sala, desde que sai da coluna de som até alcançar o microfone.

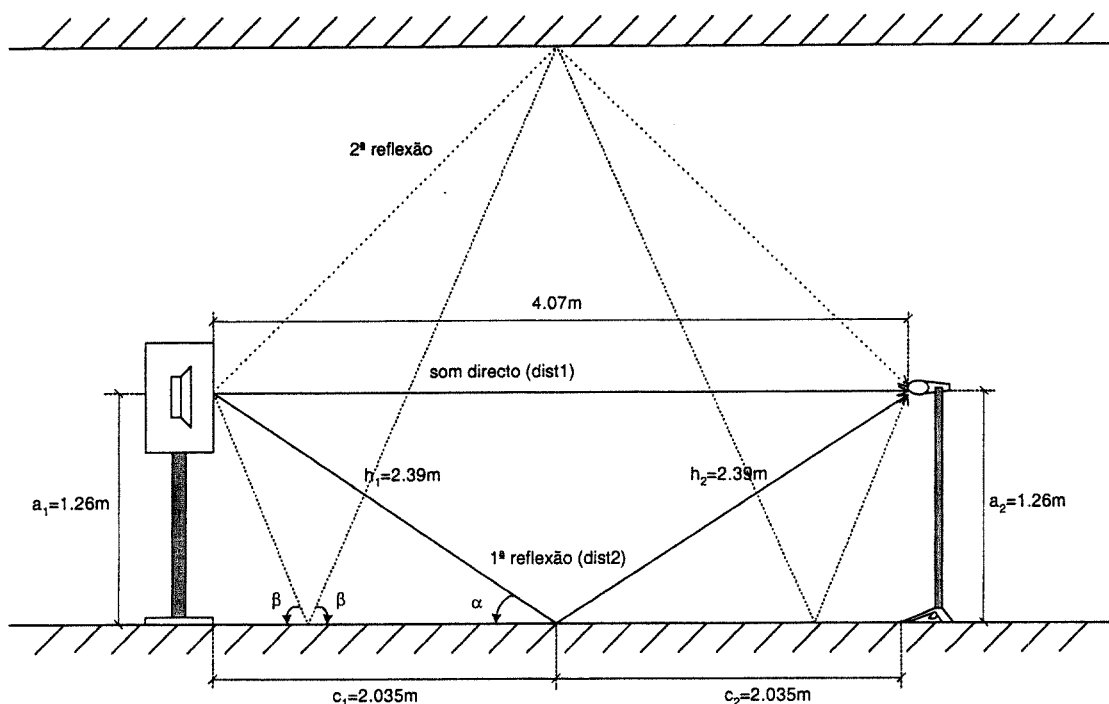


Figura 2.3 – Som directo e primeiras reflexões

Observa-se que a primeira reflexão (aquela que atinge o microfone imediatamente a seguir ao som directo), percorre a trajectória representada na Figura 2.3 por $h_1 + h_2$, ou seja, $dist2$. É de facto esta a primeira reflexão a chegar, uma vez que quer o microfone, quer a coluna se encontram bastante mais próximos do solo (a 1.26m), do que do tecto (a 1.74m) ou das paredes (das quais distam cerca de 3m). Conhecidas as dimensões a_1 , a_2 , c_1 e c_2 , foi possível pela trigonometria obter α (31.76°), logo h_1 e h_2 (2.39m), cuja soma constitui como referido a distância percorrida pela primeira reflexão, i.e., $dist2 = 4.78m$. Defina-se excesso de distância como sendo a diferença de espaço percorrido entre som directo e som reflectido. Então, o excesso de distância para a primeira reflexão é dado por $dist2 - dist1 = 0.71m$. Através das leis da física que relacionam velocidade com distância percorrida num determinado período de tempo, é possível concluir que à velocidade do som ($343ms^{-1}$), a reflexão atinge o microfone

$$\begin{aligned}
 t &= \frac{s}{v} \\
 t &= \frac{0.71}{343} \\
 t &= 2.1ms
 \end{aligned}
 \quad (2.1)$$

2.1ms após o som directo, o que de facto se verifica por inspecção da Figura 2.2 b). Não se realizaram cálculos para as reflexões seguintes a fim de não adensar o texto, mas igualmente por observação da Figura 2.2 b) constata-se que a segunda reflexão chega cerca de 4ms depois do som directo, o que parece, como esperado, indicar que esta se deve ao tecto. De facto, o excesso de distância para esta reflexão é ligeiramente inferior ao dobro do da reflexão anterior (micro e coluna encontram-se entre $\frac{1}{3}$ e $\frac{1}{2}$ da altura da sala). Assim, faz sentido que o tempo decorrido (4ms) seja também um pouco inferior ao dobro do tempo de excesso de propagação da primeira reflexão (2.1ms).

2.2. OBJECTIVOS

As reflexões verificadas na Figura 2.2 e ilustradas na Figura 2.3 são responsáveis pela alteração do som na sala, na medida em que provocam reforços e atenuações em função dos modos de ressonância da mesma que, em determinadas frequências se fazem sentir com maior intensidade, causando a irregularidade do espectro do sinal. Será o estudo dessa irregularidade espectral que permitirá ao algoritmo encontrar automática e adaptativamente um conjunto de níveis de igualização óptimo para cada sala, atendendo à sua situação particular de funcionamento, como por exemplo as suas dimensões, temperatura, humidade e quantidade de pessoas presentes dentro da mesma.

Utilizando uma técnica de filtragem adaptativa rápida no domínio das frequências, e fazendo uso de sinais excitadores possuidores de determinadas características desejadas, como por exemplo possuir uma boa SNR, pretende-se neste trabalho minimizar os efeitos acumulados condicionadores da reprodução sonora introduzidos por vários factores incluindo amplificação, transdução (i.e. características dos altifalantes) e reflexões do som nas superfícies de uma sala descritas em 2.1, através da implementação de um sistema *Real-Time Analyzer* aplicado à igualização automática.

3. CRITÉRIOS EM ACÚSTICA DE SALAS

A resposta impulsional medida entre uma fonte sonora (coluna de som) e um receptor (microfone), permite descrever as propriedades acústicas do sistema entre esses dois pontos no espaço. Refira-se que tal descrição não deve no entanto ser generalizada para representar o comportamento complexo de toda a sala.

A maioria dos parâmetros estudados em acústica pode ser extraída a partir da análise da resposta impulsional acústica do sistema, que pode por sua vez ser obtida através de diversos processos de medida.

Importa pois, fazer um estudo sobre os parâmetros que controlam a acústica de uma sala, no sentido de perceber sobre quais e de que forma será possível actuar, com o propósito de melhorar o desempenho desse mesmo sistema acústico. Este é o propósito desta secção.

3.1. MEDIÇÕES

Na presença de um sistema de reprodução ou transmissão de música ou voz em espaços fechados, a acústica deve ser prévia e cuidadosamente estudada. Na realização desse estudo são normalmente efectuadas medidas sonoras, indicadoras da resposta da sala, relativamente às quais importa considerar vários factores:

- Direccionalidade do som;
- Reflexões;
- Sensação de espaço .

A acústica de interiores determina como o som se comporta em espaços fechados, onde características físicas importantes da sala vão exercer a sua influência, pelo que importa considerar que [TUT2002]:

- Superfícies reflectivas tornam a energia do som melhor contida, i.e. aumentam a intensidade de som percebida ou *Loudness* (até certos limites);
- Diferentes gamas de frequências produzem diferentes alterações no sinal, uma vez que cada sala possui modos de reflexão próprios;
- Modos da sala poderão ser do tipo axial, tangencial e oblíquo;
- Modos causam coloração do som;
- Frequências dos modos de ressonância fazem a acústica de uma sala.

O som radiado a partir de uma fonte sonora e reflectido por uma parede rígida, forma aquilo a que se pode designar de *fonte sonora imagem*, devida à reflexão, o que pode ser entendido como se o som fosse radiado por outra fonte imaginária idêntica à fonte real e à mesma distância, colocada por trás da parede. O campo de som é, assim, formado pela componente de som directo e pelo somatório de todas as contribuições das *fontes sonoras imagem*, bem como pelo das *fontes imagem* das *imagens*.

O espalhamento geométrico de raios de som dentro de uma sala depende de diversos factores, e pode ser descrito por:

- Reflexão – processo no qual uma radiação, ao incidir numa superfície, volta para trás em vez de a penetrar;
- Absorção – efeito que se manifesta na quantidade de energia que a superfície reflectora é capaz de retirar da onda incidente;
- Refracção – desvio de direcção que os raios sofrem quando atravessam sucessivamente vários meios de densidade diferente;
- Difracção – desvio que sofrem ondas sonoras e luminosas quando passam nos bordos de um obstáculo ou através de pequenos orifícios, dependendo esse desvio do seu comprimento de onda.

No estudo de uma sala de dimensões pequenas, como é o caso da sala descrita em (2.1), importa considerar diferentes aproximações para diferentes gamas de frequência, quando se estuda o comportamento do espectro do sinal áudio, resultante dos fenómenos descritos no parágrafo anterior. Deste modo, nas situações em que o comprimento de onda é grande, comparado com as dimensões da sala, temos que:

- Não surgem ganhos devidos a ressonâncias.

Por outro lado, se o comprimento de onda do sinal é comparável às dimensões da sala, então:

- Surgem ondas estacionárias;
- Modos da sala começam a exercer a sua influência;
- Fenómenos de ressonância provocam incrementos no *loudness*.

No caso de o comprimento de onda ser muito pequeno para que a onda possa ser considerada acústica, dominam:

- Fenómenos de difusão e difracção do som.

Uma outra definição importante que importa considerar é a de eco:

- Consiste em ondas de forma idêntica, menor amplitude e atrasadas relativamente à original;
- Efeito *Haas* – o sistema auditivo integra intensidades sonoras ao longo de curtos intervalos de tempo (até cerca de 50ms), o que provoca o aumento do *Loudness*;

- Quando a réplica do sinal original dista mais de 50ms deste, então é ouvida como entidade distinta – eco.

Outro dos parâmetros importantes no estudo da acústica de salas é definido por reverberação:

- Consiste na cauda da energia sonora medida dentro de um espaço fechado, devida a múltiplas reflexões do som nas suas superfícies (componentes reflectidas chegam mais tarde);
- É boa em quantidades moderadas e má em excesso ou por total ausência, dependendo da situação. Por exemplo, a actuação de uma orquestra sinfónica gravada dentro de uma câmara anecóica soa muito mal (o som é fraco, não possui ressonância), o que dá a indicação de que a música requer alguma reverberação; se por outro lado, se pretender reproduzir a fala, então ela será mais inteligível numa sala com pouca reverberação (longa reverberação mascara as últimas consoantes da fala), o que por outras palavras significa que a fala requer pouca reverberação.

Importa igualmente considerar que em acústica de exteriores, o som pode viajar inalterado em várias direcções.

Devido a efeitos de absorção no ar e outros, a reprodução das mais altas frequências da gama audível tende a ser direcciona. Enquanto que a reprodução das frequências mais baixas tem tendência a ser omnidirecciona.

3.2. MODOS DE RESSONÂNCIA DA SALA

Atendendo às suas dimensões e para diferentes gamas de frequências, uma sala pode sofrer a influência de fenómenos ressonantes, frequências naturais e ondas estacionárias. Esses fenómenos constituem os modos da sala. Imagine-se duas paredes reflectivas de comprimento infinito com ar entre elas. Tal, pode ser considerado um sistema ressonante, com frequência de ressonância f_r calculada por:

$$f_r = \frac{c}{2L}, \quad (3.1)$$

sendo c a velocidade do som e L a distância entre as duas paredes. O sistema sofrerá também a influência de ressonâncias nas frequências $2f_r, 3f_r, \dots$, ou seja, os harmónicos de f_r .

Suponha-se agora que o sistema foi fechado e transformado numa sala rectangular. O cálculo das frequências correspondentes aos modos da sala, pode ser dado por:

$$f_r = \frac{c}{2} \sqrt{\left(\frac{p}{L}\right)^2 + \left(\frac{q}{W}\right)^2 + \left(\frac{r}{H}\right)^2}, \quad \text{com } p, q, r=0,1,2,\dots \quad (3.2)$$

onde novamente c é a velocidade do som, L é o comprimento da sala, W a largura e H a sua altura. Por p , q e r obtêm-se os harmónicos.

Os modos axiais envolvem reflexões de duas das superfícies, enquanto que os modos tangenciais estão associados a reflexões entre quatro superfícies e possuem metade da energia dos anteriores. Modos oblíquos reflectem-se por seis superfícies e possuem $\frac{1}{4}$ da energia dos modos axiais [TUT2002]. Por estes motivos, os primeiros são aqueles que exercem maior influência na resposta da sala, apesar dos modos tangencial e oblíquo também influenciarem o padrão daquela resposta. De facto, o número de modos normais de uma sala aumenta com a frequência, o que tem como consequência directa o alisamento do espectro da resposta da sala nas altas frequências [TUT2002].

Como foi já referido, o comportamento da sala é diferente, para diferentes modos de frequência. Assim, é possível afirmar que os modos decaem de forma diferente:

- O decaimento do modo depende da capacidade de absorção do material distribuído pela sala;
- A reverberação em cada banda de oitavas depende da média do decaimento dos diversos modos;
- Quanto mais alta for a frequência central da oitava, mais modos inclui;
- Cada modo ressonante tem efeito sobre uma banda estreita de frequências:
 - Esta banda de frequências é determinada pela quantidade de absorção da sala e possui cerca de 5 Hz de largura para salas típicas;
 - A largura de banda dos modos aumenta quando o tempo de reverberação diminui.

Um sinal excitador com intensidade sonora elevada, pode forçar modos adjacentes a vibrar à frequência do sinal excitador. Quando a excitação é removida, modos adjacentes decaem à sua frequência natural. Do mesmo modo, durante período de decaimento, poderão surgir trocas de energia entre modos normais [TUT2002], o que provoca a radiação de novas frequências para a sala. Dá-se nestas situações uma momentânea alteração de '*pitch*'.

3.3. LOUDNESS

Loudness é uma medida subjectiva de percepção de intensidade sonora e está relacionada com o nível de pressão sonora (SPL) que alcança a zona onde se situa o ouvinte. Depende da distância a que este se encontra da fonte, da potência desta e do coeficiente de absorção da sala. A impressão da qualidade de som numa sala é melhorada, se se garantir que ela possui suficiente *loudness*.

3.4. TEMPO DE REVERBERAÇÃO

A resposta impulsional típica de uma sala pode ser dividida em três regiões, como se ilustra na Figura 3.1. Enquanto que a primeira consiste no campo directo de som, a segunda é formada pelas primeiras reflexões do som em paredes reflectivas. Finalmente, a terceira região é formada pela parte reverberante da resposta, onde a densidade de reflexões é tão elevada, que não se distinguem reflexões individuais. São reflexões de ordem elevada em múltiplas superfícies da sala.

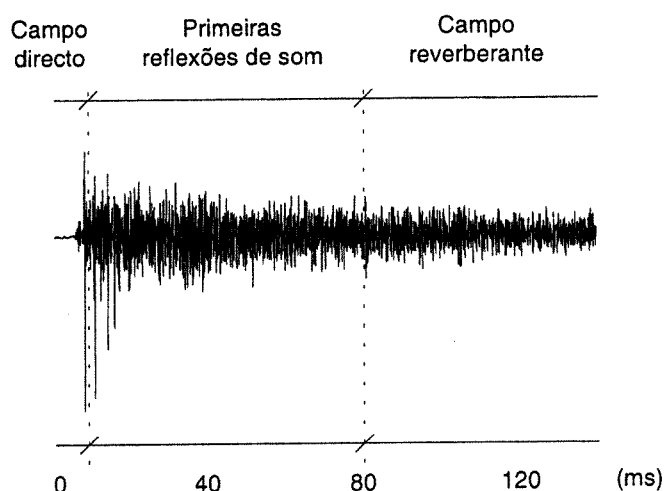


Figura 3.1 - Reverberação

A reverberação numa sala proporciona *feedback* aos ouvintes, e por esse mesmo motivo uma sala com reverberação soa muitas vezes melhor do que uma sala 'seca'. Define-se tempo de reverberação como a quantidade tempo necessária para que o nível de energia do sinal caia 60dB [Beranek90], depois da excitação ter terminado e de acordo com [Sabine1900] pode ser calculado por:

$$T_R = 0.16 \frac{V}{A}, \quad (3.3)$$

onde V representa o volume da sala e A a área total de absorção afectada pelo coeficiente de absorção. A reverberação é boa em quantidades moderadas e má em excesso. Tipicamente, o tempo de reverberação numa sala varia pouco com a frequência, e normalmente diminui com o aumento desta. A forma desta variação afecta a impressão subjectiva do som, o que pode ser exemplificado com o seguinte caso: Supondo que para uma determinada sala o tempo de reverberação para as baixas frequências é elevado, quando comparado com o das médias frequências, então o som dessa mesma sala dará a impressão de ser mais 'quente'.

Pesquisas efectuadas mostraram que as pessoas são mais influenciadas pela porção inicial de energia que alcança os ouvidos. É sabido que a percepção da direcionalidade do som é determinada apenas pelo primeiro impulso recebido, uma vez que esse é quase sempre o som directo. Impulsos secundários (reflexões) contribuem mais para a percepção da dimensão da sala. Outro efeito importante está relacionado com o período de integração. Se uma quantidade de som reflectido chegar logo após o som directo, o ouvido tem a capacidade de integrar essas

duas componentes, o que se traduz num efeito de reforço do som, contribuindo assim para a melhoria da SNR.

3.5. EDT – EARLY DECAY TIME

EDT refere-se ao intervalo de tempo que a curva do nível de energia demora a cair 10dB, depois da fonte de excitação ser abruptamente desligada. Pondera a importância subjectiva da primeira parte da reverberação. Um valor de EDT elevado indica a presença de muita *reverberação* e pouca *claridade*. As medidas de EDT possuem maior variância do que medidas do tempo de reverberação, devido a diferenças locais nas primeiras reflexões. O seu cálculo é análogo ao do tempo de reverberação, a menos dos limites de integração – que se situam agora de 0 a -10dB.

3.6. CLARIDADE

Tempos de reverberação muito longos tornam a fala difícil de compreender. A inteligibilidade do som depende da relação entre som directo e som pertencente ao campo reverberante. Salas com uma forte componente de energia nas primeiras reflexões (*early energy*) associadas a componentes tardias com pouca energia (*later energy*) possuem boa claridade. Por outro lado, claridade baixa produz a sensação de que o som não é claro, de que é excessivamente reverberante. Define-se, pois, como o logaritmo da relação entre a energia dos primeiros 80ms (para música), ou 50ms (para voz), e a restante energia do sinal:

$$C_{80} = 10 \log \frac{\int_{t=0}^{80} p^2(t) \partial t}{\int_{t=80}^{+\infty} p^2(t) \partial t} . \quad (3.4)$$

Os limites de integração encontram-se representados em milisegundos e claridade é uma grandeza expressa em dB.

3.7. DEFINIÇÃO

Consiste no logaritmo da relação entre a energia da resposta impulsional para os primeiros 80ms (no caso de música, ou 50ms no caso de voz) e a energia total da mesma resposta impulsional:

$$D_{80} = 10 \log \frac{\int_{t=0}^{80} |h(t)|^2 \partial t}{\int_{t=0}^{+\infty} |h(t)|^2 \partial t} . \quad (3.5)$$

Pode ser entendida como uma medida do grau de distinção e claridade da música ou voz radiadas numa sala.

3.8. TEMPO CENTRAL (CENTER TIME)

Equivale ao instante de tempo, na resposta impulsional, ao fim do qual se atinge 50% da energia recebida. A energia recebida para trás desse instante é igual à recebida para a frente do mesmo. Outra forma de definir este parâmetro equivale a chamar-lhe centro de gravidade da energia da resposta impulsional. Calcula-se por:

$$T_s = \frac{\int_{t=0}^{+\infty} t \times h^2(t) \partial t}{\int_{t=0}^{+\infty} h^2(t) \partial t} . \quad (3.6)$$

Um valor pequeno indica que a energia se encontra fortemente concentrada nas primeiras reflexões e por esse motivo dá a sensação de claridade, enquanto que um valor T_s elevado indica que grande parte da energia chega bastante depois do som directo, o que sugere um som reverberante.

3.9. FORÇA

Consiste na razão logarítmica entre a energia total da resposta impulsional e a energia do campo directo, capturada a uma distância de 10m da fonte.

$$G = 10 \log \frac{\int_{t=0}^{+\infty} h^2(t) \partial t}{\int_{t=0}^{+\infty} h_{a \text{ 10m da fonte}}^2(t) \partial t} \quad (3.7)$$

Descreve a influência da sala no nível de som percebido pelo ouvinte, isto é, dá uma indicação de quão difuso é o campo de som da sala.

3.10. LEF – LATERAL ENERGY FRACTION

A noção de impressão espacial refere-se à sensação de se estar envolvido pelo som, ou seja, rodeado por ele e não apenas a olhar para ele. Este parâmetro é função da contribuição relativa das reflexões laterais. Dito de outra forma, este fenómeno é determinado pelas diferenças no sinal recebido por cada ouvido. As medidas de impressão espacial são por isso calculadas com base no ângulo de cada reflexão relativa ao som directo. Pode-se desta forma afirmar que LEF é a relação entre a energia das primeiras reflexões (relativa aos primeiros 80ms) vinda das direcções laterais e entre a energia do som directo e das primeiras reflexões (primeiros 80ms) vinda de todas as direcções. Uma forma de a obter é através de:

$$LEF = \frac{\int_{t=5}^{80} h_{lateral}^2(t) \partial t}{\int_{t=0}^{80} h_{omnidir}^2(t) \partial t}, \quad (3.8)$$

onde, para se garantir a eliminação da componente de som directo no numerador da equação, se contabiliza a energia do sinal apenas após os primeiros 5ms. Assim, a energia das primeiras reflexões vinda de direcções laterais pode ser capturada usando um microfone em Figura de Oito, com o seu eixo perpendicular à direcção do campo de som directo (de modo a que a medida não venha afectada por este). Quanto à outra componente, a do denominador, relativa à energia do campo directo e das primeiras reflexões vindas de qualquer direcção, pode ser obtida através de um microfone omnidireccional muito próximo do anterior, capturando para o efeito os primeiros 80ms de sinal.

LEF é, por assim dizer, uma medida do grau de envolvimento do ouvinte no som.

As noções introduzidas nas secções anteriores são usadas na aferição da qualidade acústica em espaços fechados. Apesar de muitas dessas definições não possuírem uma relação directa com a implementação descrita nesta dissertação, importa considerar que por serem medidas comumente realizadas neste tipo de sistemas, a sua inclusão num futuro desenvolvimento deste trabalho poderá desempenhar um papel importante no equilíbrio da resposta em frequência da sala.

4. CORRECÇÃO DA RESPOSTA ACÚSTICA DA SALA

Para obter a resposta acústica de uma determinada sala, utilizam-se sinais adequados para a sua excitação, com os quais se pretende percorrer toda a gama de frequências a testar. Posteriormente, um dispositivo é responsável por capturar a resposta a essa excitação que é então comparada com o sinal original, daí se obtendo informação relativa à influência da sala. Neste capítulo são descritos vários tipos de sinais excitadores, todos eles possuidores de características distintas. É, de igual modo, apresentada a teoria relativa à aproximação de uma resposta obtida relativamente a uma resposta desejada, inserida na área do processamento adaptativo de sinais.

4.1. TIPOS DE EXCITAÇÃO DA SALA

Na excitação de uma sala, é importante utilizar sinais com elevada energia, de forma a que a relação SNR seja também elevada. Neste tipo de medidas, há um certo grau de não linearidade associado. Do mesmo modo, é típico ocorrer uma variância temporal nas medidas capturadas, que é mais evidente em alguns tipos de excitação. Comparações a propósito destes dois e de vários outros factores são discutidas nesta secção para os diferentes tipos de excitação tratados a seguir.

4.1.1. IMPULSIVA

A forma mais directa de excitar uma sala consiste em usar impulsos, idealmente de *Dirac*, capazes de excitar todas as frequências de forma linear. Deste modo, a resposta capturada não é mais do que a correspondente à resposta impulsional desejada, ou seja, um sinal de espectro plano nas frequências convoluído com a resposta impulsional da sala. Posteriormente, aplica-se-lhe uma FFT (Transformada de Fourier Rápida) para estimar a função de transferência da sala.

Contudo, impulsos deste género possuem, na prática, energia reduzida, não sendo por isso especialmente indicados para excitação de sistemas acústicos. Terão, porventura, maior aplicação no teste de sistemas electrónicos.

Uma forma de aumentar o SNR, consiste em repetir os impulsos periodicamente, e fazer a média das respostas em cada período. Deste modo, a resposta impulsional obtida é semelhante à anterior, desde que o seu comprimento seja inferior ao período em análise.

Uma técnica adicionalmente utilizada consiste em dar pré-ênfase ao sinal de excitação de modo a adaptá-lo ao ruído da sala (que se presume aditivo) e no processo de captura aplicar-lhe a curva inversa, isto é, fazer o de-ênfase, de modo a atenuar a influência do ruído da sala.

Este tipo de impulsos não permite efectuar medidas de distorção. Em contrapartida, não é tão sensível a efeitos de variações nos tempos. Não requer algoritmos de processamento de sinal

complexos, o que o torna um método interessante pela simplicidade associada [Müller&Massarini2001].

4.1.2. ALEATÓRIA

A ideia básica na utilização de uma excitação aleatória, consiste em comparar os espectros do sinal excitador (aplicado pela coluna de som à sala) com o do sinal capturado pelo microfone (que consiste no anterior após ter percorrido a sala), através da divisão do último pelo primeiro.

O caminho de transmissão acústica deve ser conhecido, de forma a fornecer um atraso ao sinal de referência, para permitir a divisão dos espectros relativos ao mesmo sinal de base.

Descontando o atraso relativo, ambos os sinais devem ser capturados e processados ao mesmo tempo. A natureza dos sinais excitadores usados neste método (ruído branco, rosa ou outros), favorece o aparecimento de um espectro muito recortado como se ilustra na Figura 4.5, onde por se usar uma estimativa de curto tempo, podem surgir vales acentuados, o que torna necessário fazer a média de vários espectros ao longo do tempo.

Para algumas frequências, a relação SNR pode ser insuficiente, e se tal acontecer, a medida à qual correspondem deverá ser retirada da média, de forma a eliminar erros no cálculo da resposta em frequência da sala.

Para atenuar fugas espectrais, a trama de sinal a analisar deverá ser janelada.

4.1.3. SINAIS MUSICAIS

Ainda baseado na técnica descrita anteriormente, existe um outro método alternativo que permite uma funcionalidade de interesse prático acrescido: a possibilidade do sinal excitador ser música. Isto permite que sejam efectuadas medidas no decorrer de um espectáculo, para correcção da acústica em tempo-real.

Contudo, a música não possui uma distribuição espectral uniforme, o que requereria períodos de cálculo de médias espectrais ainda mais longos, de modo a obter-se um resultado fiável. Por este motivo, é comum recomendar-se [Müller&Massarini2001] o uso de fontes de ruído não correlacionadas (e.g. ruído branco).

4.1.4. VARRIMENTO DISCRETO DE SINUSÓIDES (STEPPED SINE)

Uma das formas de obter a resposta impulsional de uma sala consiste em excitá-la com uma sinusóide pura, fazendo aumentar a sua frequência ao longo do tempo. O sistema receptor deverá capturar esse sinal, filtrá-lo no domínio dos tempos ou aplicar-lhe uma FFT e através desta obter no domínio das frequências, a relação de magnitude entre a sinusóide recebida e a enviada. Esta segunda hipótese revela-se superior quando comparada com filtragem no domínio dos tempos, uma vez que permite a rejeição completa de outras componentes espectrais. Aquele outro método envolveria o uso de filtros passa-banda, menos eficazes na referida rejeição de sinais indesejados (devido à falta de selectividade e precisão). No uso de

uma FFT, o comprimento da sinusóide deverá ser múltiplo inteiro do comprimento da transformada, para que não haja fugas espectrais, o que se consegue gerando (por síntese) o sinal sinusoidal excitador digitalmente. Depois de efectuada a medida para uma determinada frequência, esta é aumentada do lado excitador, multiplicando-a por um valor fixo relativamente à frequência anterior, de modo a obtermos uma representação logarítmica do espectro. A maior vantagem deste método reside na boa relação sinal ruído (SNR) que pode ser obtida numa única medida, uma vez que toda a energia se encontra concentrada numa única frequência.

Este método, é porventura, aquele que mais tempo consome na obtenção da resposta em frequência da sala. De cada vez que se altera a frequência do seno excitador, é necessário esperar até a sala atingir um estado estacionário. Só depois se fazem capturas. Quando a sala possui ressonâncias com elevado factor de qualidade (Q), esse tempo de espera deve ser elevado, de modo a obter-se um erro desprezável. Adicionalmente, a diferença de tempo entre som directo e som reflectido deverá ser superior ao da janela de captura².

É, no entanto, um método muito usado em medidas de precisão, na calibração de material electrónico e em medidas de distorção (há a facilidade de detectar harmónicos com grande precisão através da FFT).

4.1.5. SEQUÊNCIAS DE COMPRIMENTO MÁXIMO (MLS)

Sinais MLS (*Maximum Length Sequences*) são sequências pseudo-aleatórias que proporcionam outra forma de excitar a acústica de uma sala. São na realidade sequências binárias aleatórias, de repetitividade (periodicidade) conhecida e controlada através do correcto dimensionamento do 'shift register' e da localização das entradas da 'gate' XOR que formam o algoritmo gerador do sinal. Um algoritmo de N andares será capaz de gerar $2^N - 1$ estados diferentes, todos, excepto o estado 'tudo zeros', antes que o sinal se repita. Com $N=17$ andares, um algoritmo deste género é capaz de gerar 131072 estados até se repetir, o que a uma frequência de amostragem de 48kHz nos permite excitar a sala com um sinal de cerca de 2.7 segundos. Esta duração é suficiente para excitar todos os modos da sala e esperar que estes atinjam um regime de funcionamento estacionário, para depois se proceder à captura estável da resposta.

Tipicamente, são necessárias grandes quantidades de memória para armazenar o sinal excitador e posteriormente a resposta da sala.

Há alguns anos atrás, as limitações de memória em sistemas computacionais eram contornadas gerando estes sinais através de *hardware*. Com a evolução da indústria de semicondutores e, mais especificamente, com o surgimento no mercado de DSPs rápidos com grande quantidade de memória disponível a custo reduzido, tornou-se viável e até mais eficiente, gerar este tipo de sinais por *software* e enviá-los para a saída via conversor D/A, posteriormente amplificando-os e excitando a sala com eles.

² Em salas pequenas as primeiras reflexões podem eventualmente surgir poucos milissegundos após o som directo, e nesse caso a diferença entre aqueles dois tempos é inferior ao da janela de captura. A Figura 2.2 b) dá indicação das primeiras reflexões para a sala em estudo neste trabalho e é um exemplo do que acaba de ser descrito.

Uma das características que torna as sequências MLS desejadas no cálculo da resposta impulsional de uma sala reside no facto da sua autocorrelação se aproximar de um impulso unitário discreto, o que indica um espectro branco. Comparando este método (excitação por sinais MLS) com o referido em 4.1.1 (excitação impulsiva), observa-se que para amplitudes de sinal idênticas se consegue excitar o DUT com uma quantidade de energia superior, na medida em que o sinal excitador é ‘esticado’ para durar o tempo de captura da medida. Tal facto traduz-se numa melhoria da relação sinal ruído (SNR).

Como desvantagens, deve-se referir que este método apresenta uma grande vulnerabilidade à distorção e a efeitos de variação temporal.

4.2. FILTRAGEM ADAPTATIVA

Em certas aplicações na área da acústica, como por exemplo no cancelamento de eco, o filtro adaptativo necessita de uma resposta impulsional longa (i.e., de memória longa), para poder lidar com uma igualmente longa duração do eco. Usando algoritmos LMS (*Least Mean Square standard*), em que a adaptação é feita amostra a amostra no domínio temporal, observa-se que a complexidade computacional de um filtro deste género é elevada. Uma forma de contornar este problema, consiste em utilizar filtragem adaptativa no domínio das frequências (FDAF). Uma vez que existem algoritmos que permitem transpor a representação temporal de um sinal para o domínio das frequências, permitindo de seguida a operação inversa, com reconstrução perfeita do sinal, é possível realizar a filtragem, bem como a adaptação dos parâmetros do filtro, no domínio das frequências. Esta técnica, designada por filtragem adaptativa no domínio das frequências (FDAF), proporciona um caminho alternativo ao método LMS no domínio dos tempos. O algoritmo de LMS rápido, assim chamado por proporcionar um algoritmo eficiente do ponto de vista computacional para a realização de filtragem FIR adaptativa com memória longa, é baseado na filtragem adaptativa por blocos. Este algoritmo explora a vantagem oferecida pelo método de convolução rápida no domínio das frequências, conhecido por *Overlap-and-Add (OA)*, ou *Overlap-and-Save (OS)*, que é implementado usando algoritmos de FFT (*Fast Fourier Transform*). O algoritmo de LMS rápido, exhibe propriedades de convergência parecidas com as do algoritmo LMS *standard*, na medida em que ambos tentam minimizar a mesma função de erro quadrático médio. As diferenças residem no facto de o primeiro algoritmo ter um limite de estabilidade mais apertado do que o segundo, proporcionando no entanto, uma estimativa do gradiente mais precisa, precisão essa que aumenta com o tamanho do bloco. Esta melhoria não implica no entanto, uma convergência mais rápida [Haykin96].

Assim, em resultado dos argumentos anteriores e face às características do sistema desenvolvido e aqui descrito, na realização deste trabalho faz-se uso de filtragem no domínio das frequências.

4.2.1. EXTRACÇÃO DE CARACTERÍSTICAS

Com base nos métodos descritos anteriormente na secção 4.1, foi-nos possível seleccionar algoritmos para implementação em DSP e em tempo-real, que nos permitissem obter

conclusões quanto à adaptatividade e correcção automática da acústica de uma sala por igualização automática. Os métodos testados usam excitação por ruído aleatório branco descrito em 4.1.2 e através de sequências MLS indicadas em 4.1.5.

Uma vez excitada a sala (DUT) com um dos sinais descritos $x(t)$, a sua resposta impulsional $h(t)$ é convoluída com este e posteriormente captada por um microfone de medida (de precisão) em $y(t)$.

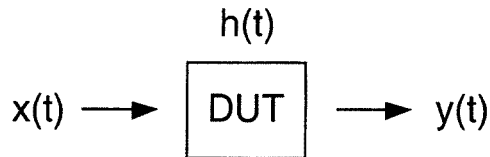


Figura 4.1 - Sistema Linear e Invariante nos Tempos

O processo descrito na Figura 4.1 diz-nos que uma determinada saída $y(t)$ reage a uma entrada arbitrária $x(t)$, com uma resposta que depende dessa mesma entrada $x(t)$ e da resposta impulsional do sistema $h(t)$.

Fazendo a analogia para um sistema discreto nos tempos, é possível representar a Figura 4.1 através das equações (4.1) e (4.2), onde o símbolo $*$ é utilizado para indicar a operação de convolução de sinais discretos.

$$y(n) = x(n) * h(n) \quad (4.1)$$

$$y(n) = \sum_{k=-\infty}^{+\infty} x(k)h(n-k) \quad (4.2)$$

A resposta impulsional acústica descreve o comportamento entre dois pontos no espaço, de um sistema acústico que se admite poder ser aproximado por um sistema linear e invariante nos tempos (LTI)³. Daqui se infere que o sistema em análise pode ser descrito como sendo LTI. Através do estudo da resposta impulsional do sistema, é possível extrair características que permitam actuar, de forma a corrigir perturbações provocadas pela sala.

4.2.2. MÉTODO LMS STANDARD

Com base na Figura 4.1 e nas equações (4.1) e (4.2) definiu-se um sistema LTI e suas propriedades. Vamos agora supor que a resposta impulsional $h(n)$ se torna adaptável em função de um determinado critério, que no caso será o de minimização de um erro $e(n)$. Ilustra-se na Figura 4.2 um sistema deste tipo.

³ Um sistema diz-se linear se for aditivo e homogéneo. Se produz, respectivamente, para duas entradas x_1 e x_2 as saídas y_1 e y_2 , entende-se que ele é aditivo se para a entrada x_1+x_2 produzir a saída y_1+y_2 . Do mesmo modo, diz-se homogéneo se a entrada $c.x_1$ resultar na saída $c.y_1$, sendo c uma constante qualquer. Por outro lado, diz-se que um sistema é invariante nos tempos, se para qualquer instante de tempo a sua resposta impulsional $h(t)$ se mantiver constante.

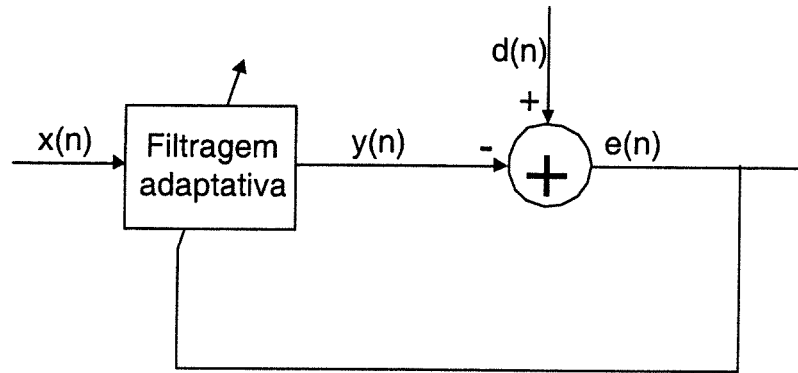


Figura 4.2 - Filtragem adaptativa

Filtros Wiener – filtros óptimos

Sendo o sistema $h(n)$ (variável) excitado por uma entrada $x(n)$, produz na saída uma resposta $y(n)$. Com a utilização deste tipo de filtragem adaptativa pretende-se conceber um sistema cuja saída $y(n)$ produza uma estimativa da resposta desejada $d(n)$, de tal forma que o erro quadrático médio esperado $E\{|e(n)|^2\}$, sendo $e(n)$ definido como a diferença entre a resposta desejada $d(n)$ e a resposta actual $y(n)$, seja tão pequeno quanto possível. A aproximação de $y(n)$ a $d(n)$ faz-se em iterações sucessivas actuando por realimentação em $h(n)$, com uma velocidade de convergência controlada. Um filtro deste género possui restrições e deverá ser:

1. Linear, o que facilita o uso de matemática;
2. Discreto, para que seja implementável usando tecnologia digital.

Ambas as restrições são respeitadas, como descrito anteriormente.

Duas outras escolhas deverão ser tomadas, no que diz respeito à especificação do filtro:

1. A resposta impulsional do filtro poderá ter duração finita ou infinita;
2. Tipo de critério estatístico utilizado para optimização deverá ser definido.

O ponto 1. foi já esclarecido anteriormente, uma vez que se tratam de filtros FIR, logo, de resposta de duração finita. Esta escolha é particularmente conveniente dado que um filtro FIR é intrinsecamente estável. O critério de convergência utilizado consiste na minimização do erro quadrático médio de $e(n)$. Na realidade, procura-se otimizar o desenho do filtro, minimizando uma função de custo J do erro quadrático médio. Essa função de custo J tem um mínimo distinto que define de forma única a implementação estatística óptima do filtro [Haykin96].

Princípio da Ortogonalidade

Considere-se de novo a Figura 4.2, onde a entrada do filtro é indicada pela sequência $x(n)$ e a sua resposta impulsional por $h(n)$, admitindo-se que é causal e de comprimento N . A saída do filtro para um instante discreto n é dada por [Oppenheim&Schafer99]:

$$y(n) = \sum_{k=0}^{N-1} h^*(k)x(n-k). \quad (4.3)$$

Assume-se estacionariedade em sentido lato dos processos estocásticos representados por $d(n)$ – sinal desejado, e $x(n)$ – sinal de entrada. A aproximação a $d(n)$ é acompanhada pelo cálculo de um erro, dado por:

$$e(n) = d(n) - y(n). \quad (4.4)$$

Para conceber o filtro optimizado, como referido anteriormente, o objectivo é minimizar o erro quadrático médio estimado através da função de custo J :

$$\begin{aligned} J &= E\{e(n)e^*(n)\} \\ &= E\{|e(n)|^2\} \end{aligned} \quad (4.5)$$

Este princípio mostra que, através da derivada de primeira ordem de J em função dos coeficientes $h(k)$, definida por $\frac{\partial J}{\partial h(k)}$, é possível calcular o valor mínimo para a função de custo J . Como demonstrado por [Haykin96], o cálculo desse gradiente conduz à expressão:

$$E\{x(n-k)e_0^*(n)\} = 0, \quad \text{com } k=0,1,2,\dots \quad (4.6)$$

onde e_0 indica o erro estimado que resulta quando o filtro opera na sua condição óptima. Assim, a condição necessária e suficiente para que a função de custo J atinja o seu valor mínimo, é que o valor do erro de estimação $e_0(n)$ seja ortogonal à amostra de entrada que entra no cálculo da estimação da resposta desejada no instante n .

O corolário do princípio de ortogonalidade pode ser obtido por análise da correlação entre $y(n)$ e o erro estimado $e(n)$. Através de (4.3) essa correlação pode ser expressa por:

$$\begin{aligned} E\{y(n)e^*(n)\} &= E\left\{\sum_{k=0}^{N-1} h^*(k)x(n-k)e^*(n)\right\} \\ &= \sum_{k=0}^{N-1} h^*(k)E\{x(n-k)e^*(n)\} \end{aligned} \quad (4.7)$$

Sabendo que e_0 é o erro estimado que resulta quando o filtro opera na sua condição óptima, podemos indicar y_0 como a saída do filtro optimizado em função do erro quadrático médio, o que, através do princípio da ortogonalidade descrito em (4.6), conduz ao resultado:

$$E\{y_0(n)e_0^*(n)\}=0 \quad (4.8)$$

Assim, quando o filtro opera na sua condição óptima, a estimativa da resposta desejada definida pela saída do filtro $y_0(n)$ e o erro de estimação $e_0(n)$ são ortogonais.

Equações Wiener-Hopf

Recordando a Figura 4.2 e assumindo que lidamos com sinais reais e filtro FIR de coeficientes reais, temos através das equações (4.3) e (4.4):

$$e(n) = d(n) - \sum_{\ell=0}^{N-1} h(\ell)x(n-\ell), \quad (4.9)$$

o que permite decompor a função de custo $J = E\{|e(n)|^2\}$ em:

$$J = E\left\{\left[d(n) - \sum_{\ell_a}^{N-1} h(\ell_a)x(n-\ell_a)\right] \cdot \left[d(n) - \sum_{\ell_b}^{N-1} h(\ell_b)x(n-\ell_b)\right]\right\} \quad (4.10)$$

e

$$\begin{aligned} J = E\left\{ \left| d(n) \right|^2 - \sum_{\ell_b}^{N-1} h(\ell_b)d(n)x(n-\ell_b) - \sum_{\ell_a}^{N-1} h(\ell_a)d(n)x(n-\ell_a) \right. \\ \left. + \sum_{\ell_a}^{N-1} \sum_{\ell_b}^{N-1} h(\ell_a)h(\ell_b)x(n-\ell_a)x(n-\ell_b) \right\} \end{aligned} \quad (4.11)$$

onde:

$$\begin{aligned} J = E\left\{ \left| d(n) \right|^2 \right\} - \sum_{\ell_b}^{N-1} h(\ell_b)E\{d(n)x(n-\ell_b)\} - \sum_{\ell_a}^{N-1} h(\ell_a)E\{d(n)x(n-\ell_a)\} \\ + \sum_{\ell_a}^{N-1} \sum_{\ell_b}^{N-1} h(\ell_a)h(\ell_b)E\{x(n-\ell_a)x(n-\ell_b)\}. \end{aligned} \quad (4.12)$$

Uma vez mais, assumindo estacionariedade em sentido lato dos processos estocásticos representados por $d(n)$ – sinal desejado, e por $x(n)$ – entrada do filtro adaptativo, é possível decompor os valores esperados para o segundo termo de (4.12) nas seguintes parcelas:

$$E\{|d(n)|^2\} = \sigma_d^2, \quad (4.13)$$

onde σ_d^2 representa a variância da resposta desejada $d(n)$ (que se presume de média nula). Da mesma forma, o segundo valor esperado de (4.12) pode ser representado da forma:

$$E\{d(n)x(n-\ell_b)\} = p(-\ell_b), \quad (4.14)$$

onde $p(-\ell_b)$ representa a função de correlação cruzada entre os sinais $d(n)$ e $x(n)$. Do mesmo modo, o termo:

$$E\{d(n)x(n-\ell_a)\} = p(-\ell_a) \quad (4.15)$$

representa igualmente correlação cruzada entre os dois sinais.

O quarto valor esperado da equação pode ser rescrito da forma:

$$\begin{aligned} E\{x(n-\ell_a)x(n-\ell_b)\} &= E\{x(m)x[m+(\ell_a-\ell_b)]\} \\ &= r(\ell_a-\ell_b) \end{aligned} \quad , \text{ com } m=n-\ell_a \quad (4.16)$$

onde $r(\ell_a-\ell_b)$ equivale à função de autocorrelação da entrada do filtro $x(n)$ para um atraso de $\ell_a-\ell_b$. Assim, (4.12) pode ser totalmente rescrita na forma:

$$J = \sigma_d^2 - \sum_{\ell_b}^{N-1} h(\ell_b)p(-\ell_b) - \sum_{\ell_a}^{N-1} h(\ell_a)p(-\ell_a) + \sum_{\ell_a}^{N-1} \sum_{\ell_b}^{N-1} h(\ell_a)h(\ell_b)r(\ell_a-\ell_b), \quad (4.17)$$

equação esta que mostra que, sendo o sinal de entrada e a resposta desejada estacionários, então a função de custo J , ou erro quadrático médio, é uma função de segunda ordem dos coeficientes do filtro. Prova-se [Haykin96] que essa relação de J em função dos coeficientes corresponde a uma parabolóide, caracterizada por um mínimo único. É a chamada *superfície de performance de erro* do filtro. No ponto onde a função J atinge o seu mínimo, o seu gradiente em função dos coeficientes do filtro anula-se. A Figura 4.3 dá um exemplo da função de custo J para um filtro de dois coeficientes, caso ilustrável de uma superfície do género, já que é tridimensional. A dimensão de uma função deste tipo é de ordem $N+1$, onde N representa o número de coeficientes do filtro.

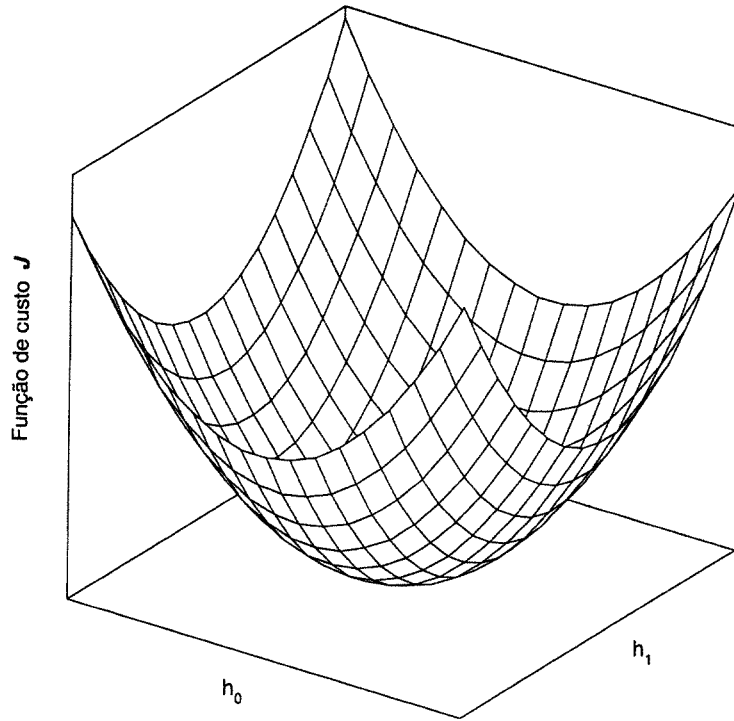


Figura 4.3 - Performance da superfície de erro para filtro de dois coeficientes h_0 e h_1

Torna-se deste modo necessário estudar o comportamento do gradiente da função de custo em função dos coeficientes, com vista à obtenção do ponto óptimo de funcionamento do filtro. Derivando J em ordem aos coeficientes, vem:

$$\begin{aligned} \frac{\partial J}{\partial h(k)} = & -E\{d(n)x(n-k)\} - E\{d(n)x(n-k)\} + \sum_{\ell=0}^{N-1} h(\ell)E\{x(n-k)x(n-\ell)\} \\ & + \sum_{\ell=0}^{N-1} h(\ell)E\{x(n-\ell)x(n-k)\}, \end{aligned} \quad (4.18)$$

que simplificando dá:

$$\frac{\partial J}{\partial h(k)} = -2E\{d(n)x(n-k)\} + 2 \sum_{\ell=0}^{N-1} h(\ell)E\{x(n-\ell)x(n-k)\}, \quad (4.19)$$

o que a partir de (4.15) e (4.16) pode ser rescrito como:

$$\frac{\partial J}{\partial h(k)} = -2p(-k) + 2 \sum_{\ell=0}^{N-1} h(\ell)r(\ell-k). \quad (4.20)$$

Calculando a expressão para o ponto óptimo do filtro, em que a derivada da superfície de erro se anula:

$$\frac{\partial J}{\partial h(k)} = 0, \quad \text{com } k=0,1,2,\dots,N-1 \quad (4.21)$$

temos que:

$$p(-k) = \sum_{\ell=0}^{N-1} h(\ell)r(\ell-k), \quad (4.22)$$

ou seja, para o referido ponto, o filtro opera na sua condição óptima se os coeficientes óptimos $h(\ell)$ forem obtidos pelo sistema de equações de Wiener-Hopf em (4.22), que definem os coeficientes óptimos do filtro em termos da função de autocorrelação da entrada do filtro e da correlação cruzada entre a entrada e a resposta desejada [Haykin96].

Algoritmo de declive mais negativo (*Steepest Descent*)

A análise anterior permitiu encontrar o vector gradiente para o instante n . Desconhecendo-se as estatísticas $h(\ell)$ e $p(\ell)$ e de modo a procurar um critério de adaptação iterativa do filtro FIR usado aproximando-o do óptimo, parte-se tipicamente da solução nula ($h(k)=0, \forall k$) e avalia-se para cada instante n , a direcção mais negativa do gradiente da função J , em ordem aos coeficientes do filtro, que é usada para forçar a convergência destes coeficientes para os valores óptimos.

Se $h(k)$ for o vector de coeficientes para o instante n , o novo vector de coeficientes actualizado para o instante $n+1$ será:

$$h_{n+1}(k) = h_n(k) - \frac{1}{2} \mu \frac{\partial J}{\partial h_n(k)}, \quad (4.23)$$

onde μ define o passo de adaptação, ou seja, a velocidade de convergência do algoritmo. Substituindo (4.19) em (4.23), observa-se que:

$$h_{n+1}(k) = h_n(k) + \mu \left[E\{d(n)x(n-k)\} - \sum_{\ell=0}^{N-1} h_n(\ell) E\{x(n-\ell)x(n-k)\} \right], \quad (4.24)$$

ou,

$$h_{n+1}(k) = h_n(k) + \mu E \left\{ x(n-k) \left[d(n) - \sum_{\ell=0}^{N-1} h_n(\ell) x(n-\ell) \right] \right\}. \quad (4.25)$$

Substituindo a última parte do segundo termo de (4.25) por (4.9), vem:

$$h_{n+1}(k) = h_n(k) + \mu E \{ x(n-k)e(n) \} \quad (4.26)$$

Esta solução corresponde ao algoritmo de '*Steepest Descent*' que usa um gradiente determinístico no cálculo recursivo dos coeficientes do filtro de Wiener.

Algoritmo LMS

Um algoritmo LMS é um algoritmo de gradiente estocástico, que se distingue do anterior por não requerer medidas de médias estatísticas. Neste sentido, a simplicidade tornou o LMS o algoritmo *standard*, no que respeita a filtragem adaptativa. Os valores obtidos no cálculo dos coeficientes através deste algoritmo, representam uma estimativa instantânea à solução $h(n)$ de Wiener e são por isso designados por $\hat{h}(n)$. O algoritmo LMS usa o produto $x(n-k)e(n)$ como estimativa do elemento k no vector gradiente e remove o operador E (espectância). Por este motivo, o cálculo recursivo dos coeficientes neste tipo de algoritmo, sofre do fenómeno designado por *ruído de gradiente*. A equação (4.26) é então rescrita da forma:

$$\hat{h}(n+1) = \hat{h}(n) + \mu x(n-k)e(n) \quad (4.27)$$

Pelo método de '*Steepest Descent*', foi demonstrado [Haykin96], que o algoritmo calculava adaptativamente novos coeficientes $h(n)$ descendo pela *superfície de performance de erro*, acompanhando uma trajectória determinística que terminava na solução de Wiener. O algoritmo LMS comporta-se de forma ligeiramente diferente, devido à presença do referido *ruído de gradiente*. Ao invés de terminar na solução de Wiener, o algoritmo executa um movimento aleatório em torno do ponto mínimo da *superfície de performance de erro*. A diferença entre o valor $J(\infty)$ encontrado por LMS, e o valor J_{min} para a solução óptima de Wiener, é o preço a pagar pela utilização de um método estocástico adaptativo, em vez de uma aproximação determinística usando o método de '*Steepest Descent*'. Esta medida de desajuste entre os dois métodos é, no entanto, controlável através de uma escolha adequada da constante μ .

4.2.3. ANALOGIA ENTRE ALGORITMO IMPLEMENTADO E LMS STANDARD

A filtragem adaptativa pode ser dividida em vários tipos de modelos. A principal diferença entre eles reside na forma como a resposta desejada é extraída. Distinguem-se quatro categorias de aplicações em filtragem adaptativa:

1. Identificação – extracção de modelos matemáticos sobre sistemas desconhecidos;
2. Modelização inversa – deseja-se encontrar o melhor modelo que se adapte ao sistema no sentido de anular o efeito deste.
3. Predição – o filtro adaptativo deverá fornecer a melhor predição do valor de um sinal aleatório.
4. Cancelamento de interferências – o filtro adaptativo é usado para cancelar interferências desconhecidas contidas num sinal primário.

É comum ver este último tipo de modelo implementado na remoção de componentes espectrais indesejados, como sejam os 50Hz da rede, ou no cancelamento de ecos e *feedback*.

O sistema implementado na realização deste trabalho não faz adaptação no sentido convencional como usualmente caracterizada na literatura. Por outras palavras, os coeficientes do filtro não serão ‘taps’ pesados para convolução directa com o sinal a corrigir no domínio dos tempos, mas sim incrementos ou decrementos nos níveis (parâmetros) de um equalizador gráfico digital de 20 bandas. Os coeficientes agora definidos por $w(k)$ são, pois, grandezas no domínio das frequências e o tipo de filtragem adaptativa implementada neste trabalho pode ser entendido como a mistura de um Modelo de Identificação com um outro de Modelização Inversa. Dito de outra forma, a saída do filtro $y(k)$ não é agora uma grandeza temporal, mas sim espectral. O índice k em $y(k)$ representa uma das 20 bandas a serem equalizadas. Do mesmo modo, a resposta desejada $d(k)$ é também uma grandeza no domínio das frequências. Consiste no espectro desejado que, na aplicação em causa, se quer ‘flat’. Contudo, $d(k)$ pode ter uma especificação diversa. O erro resultante da diferença daquelas duas grandezas – $e(k)$, é também ele uma grandeza nas frequências. $f(k)$ representa a resposta da sala a uma excitação sonora e G identifica o volume (em dBs e constante a todas as frequências), às variações do qual o sistema deverá ser imune (dentro de certos limites para que se salvguarde a SNR e evite distorções). Uma descrição possível do sistema encontra-se ilustrada na Figura 4.4.

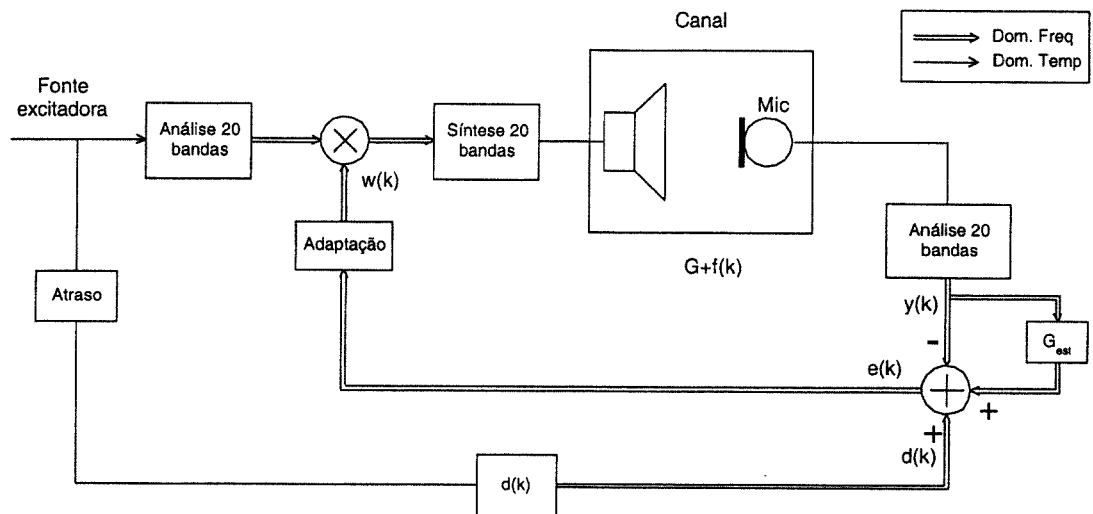


Figura 4.4 - Modelo da filtragem adaptativa usada na realização do trabalho

Note-se que uma vez que as grandezas em causa são trabalhadas em escala logarítmica, todas as influências dos sucessivos blocos descritos na figura são tratadas como somas, em vez de multiplicações. Também deve ser referido que o espectro desejado $d(k)$ deveria ser tratado de forma idêntica a $y(k)$, o que equivale a dizer que também uma constante D_{est} (fruto do valor médio de $d(k)$), deveria ser subtraída àquele valor. Tal não acontece apenas porque no caso da aplicação em causa, $d(k)$ assume para todas as bandas o valor constante de 0dB, o que constitui um espectro de referência constante ou ‘flat’, como pretendido.

Tendo sempre em consideração a analogia com o algoritmo LMS *standard*, procede-se à implementação de um outro algoritmo que vai de igual modo efectuar uma correcção, por

actualizações sucessivas dos coeficientes $w(k)$, calculadas com base na medida de um erro $e(k)$, fruto da diferença entre uma determinada resposta desejada $d(k)$ e a saída do filtro $y(k)$, segundo a equação:

$$e(k) = d(k) - [y(k) - G_{est}], \quad (4.28)$$

ou:

$$e(k) = d(k) - [w(k) + f(k) + G - G_{est}]. \quad (4.29)$$

G_{est} representa uma medida simples do valor médio da gama dinâmica do sinal $y(k)$ e a sua inclusão na fórmula prende-se com a tentativa de tornar o sistema imune a variações no ganho G da sala:

$$G_{est} = \frac{MAX[y(k)] + MIN[y(k)]}{2}. \quad (4.30)$$

Para simplificar, (4.29) pode então ser escrita de outra forma:

$$e(k) = d(k) - [w(k) + f(k)]. \quad (4.31)$$

Note-se que (4.31) não considera o efeito que a influência do erro existente nas outras bandas $\ell \neq k$ possa ter sobre a banda k . Se tal acontecesse, (4.31) deveria ser rescrita da seguinte forma:

$$e(k) = d(k) - [w(k) + f(k)] + \sum_{\ell \neq k} \xi(\ell), \text{ com } k=1, \dots, 20 \quad (4.32)$$

onde $\xi(\ell)$ representa a referida influência do erro das outras bandas $\ell \neq k$ na banda k . O erro $\xi(\ell)$ depende da selectividade espectral do banco de filtros utilizado, e será tanto menor quanto maior for aquela. Como no caso da aplicação desenvolvida, a selectividade obtida pelo método usado de filtragem no domínio das frequências é elevada, presume-se ortogonalidade entre bandas e despreza-se $\xi(\ell)$. Assim, o erro vem representado por (4.31), como inicialmente indicado.

Assuma-se de novo uma função de custo J , mantendo presente que o objectivo é minimizar o erro quadrático médio estimado. Uma vez que todas as bandas contribuem para o cálculo do erro, a função J vem:

$$J = E \left\{ \sum_{\ell=1}^{20} |e(\ell)|^2 \right\} \quad (4.33)$$

Substituindo por (4.31):

$$J = E \left\{ \sum_{\ell=1}^{20} [d(\ell) - w(\ell) - f(\ell)][d(\ell) - w(\ell) - f(\ell)] \right\}, \quad (4.34)$$

ou:

$$\begin{aligned}
 J = E\{ \sum_{\ell} d^2(\ell) - \sum_{\ell} d(\ell)w(\ell) - \sum_{\ell} d(\ell)f(\ell) \\
 - \sum_{\ell} w(\ell)d(\ell) + \sum_{\ell} w^2(\ell) + \sum_{\ell} w(\ell)f(\ell) \\
 - \sum_{\ell} f(\ell)d(\ell) + \sum_{\ell} f(\ell)w(\ell) + \sum_{\ell} f^2(\ell) \}.
 \end{aligned} \tag{4.35}$$

Agrupando, obtém-se:

$$J = \sum_{\ell} E\{d^2(\ell)\} - 2\sum_{\ell} w(\ell)E\{d(\ell)\} - 2\sum_{\ell} E\{d(\ell)f(\ell)\} + \sum_{\ell} w^2(\ell) + 2\sum_{\ell} w(\ell)E\{f(\ell)\} + \sum_{\ell} E\{f^2(\ell)\}. \tag{4.36}$$

O gradiente da função de custo J em função dos coeficientes do filtro anula-se, quando o filtro atinge o seu ponto óptimo de funcionamento. É esse ponto óptimo que se deseja encontrar, calculando o gradiente:

$$\frac{\partial J}{\partial w(k)} = -2E\{d(k)\} + 2w(k) + 2E\{f(k)\}. \tag{4.37}$$

Fazendo a analogia com (4.23):

$$w_{n+1}(k) = w_n(k) - \frac{1}{2}\mu \frac{\partial J}{\partial w_n(k)}, \tag{4.38}$$

ou,

$$w_{n+1}(k) = w_n(k) + \mu [E\{d(k)\} - w_n(k) - E\{f(k)\}], \tag{4.39}$$

e, fazendo novamente a analogia com o método LMS, obtém-se:

$$w_{n+1}(k) = w_n(k) + \mu [d(k) - E\{w_n(k) + f(k)\}] \tag{4.40}$$

em que se presume que a resposta desejada é constante. Este resultado, juntamente com (4.28) e (4.29) permite, como pretendido, concluir que:

$$w_{n+1}(k) = w_n(k) + \mu [d(k) - [E\{y(k)\} - E\{G_{est}\}]], \tag{4.41}$$

ou seja, com base na extracção do valor médio G_{est} do esperado de $y(k)$, obtém-se um sinal que subtraído ao desejado 'flat' $d(k)$, fornece a indicação do gradiente com que se procede à alteração dos coeficientes antigos $w_n(k)$ para obtenção dos novos $w_{n+1}(k)$. De realçar que o objectivo deste trabalho prende-se com o alisamento do espectro do sinal, isto é, com um sinal desejado $d(k)$ 'flat' que por se tratar de uma grandeza logarítmica, apresenta contribuição nula (ajuste para um objectivo de '0 dB'), o que nos permite simplificar (4.41) na forma:

$$w_{n+1}(k) = w_n(k) - \mu[E\{y(k)\} - G_{est}] \quad (4.42)$$

A indicação de um valor esperado $E\{y(k)\}$ para a resposta da sala capturada, parece sugerir que se proceda a uma média de vários espectros, em vez de se processar apenas uma trama em cada ajuste iterativo do algoritmo. Se fosse este o caso, (4.42) poderia ser ainda mais simplificada e consistiria apenas em:

$$w_{n+1}(k) = w_n(k) - \mu[y(k) - G_{est}], \quad (4.43)$$

o que é um resultado intuitivo.

Este trabalho foi implementado com base em (4.42), uma vez que se efectua a captura e média, como será descrito adiante neste texto, de 50 tramas do espectro $y(k)$ para a realização de cada iteração.

4.3. CÁLCULO ADAPTATIVO DE NOVOS PARÂMETROS DE IGUALIZAÇÃO

4.3.1. EXCITAÇÃO IMPULSIONAL POR DISPARO DE PISTOLA SONORA

A resposta impulsional unitária é definida como sendo a resposta de um sistema a um impulso

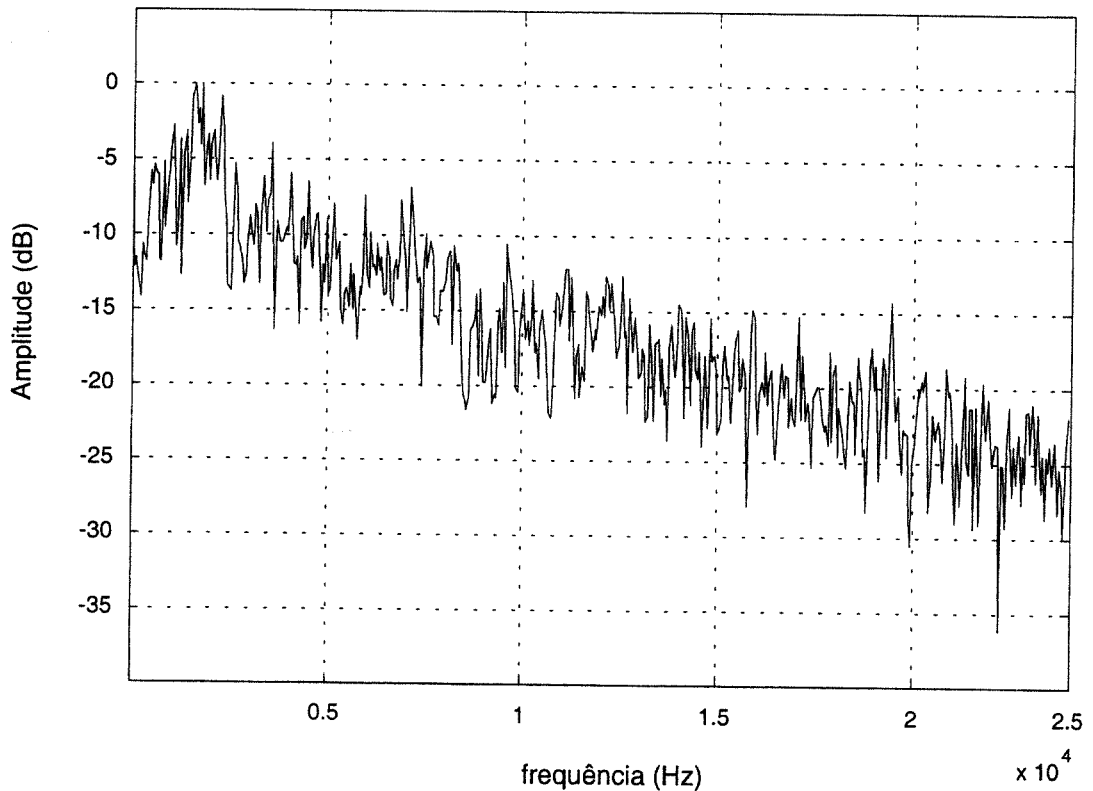


Figura 4.5 – Resposta da sala a um disparo de pistola

ideal, i.e. um impulso unitário. Esta resposta impulsional $h(n)$ descreve o comportamento do sistema para qualquer resposta. Através da análise de (4.1) observa-se que para uma entrada $x(n)=\delta(n)$, a saída do sistema será a sua resposta, porque $y(n)=\delta(n)*h(n)=h(n)$.

Possuindo uma plataforma DSP e um sistema de amplificação, reprodução e captação de som (como é o caso), criar um impulso de *Dirac* capaz de excitar sonoramente uma sala é uma tarefa relativamente fácil dos pontos de vista algorítmico e de processamento de sinal. O problema na realização desta tarefa prende-se com o facto da energia contida num sinal deste tipo ser bastante reduzida. Por este motivo, uma das formas tipicamente usadas no estudo da acústica de espaços, consiste na utilização de uma pistola (sonora) para excitação da sala. No âmbito desta dissertação, realizaram-se ensaios utilizando esta técnica. O objectivo consiste na caracterização da resposta acústica da sala, que pode ser observada por inspecção da Figura 4.5. É possível verificar que a resposta obtida, está bastante próxima daquela que será observada com os métodos automáticos de excitação descritos mais à frente neste texto. Contudo, o uso de pistolas sonoras não é recomendado, uma vez que possuem péssima repetitividade e direcionalidade imprevisível [Müller&Massarini2001].

Quer-se, pois, um método de excitação automático, do ponto de vista de excitação, captura e adaptação da resposta da sala. Os próximos capítulos descrevem sistemas deste género.

4.3.2. ALGORITMO COM EXCITAÇÃO POR RUÍDO BRANCO

O primeiro método implementado usa ruído branco como fonte sonora excitadora do sistema. Sejam $X(f)$, $H(f)$ e $Y(f)$ as representações em frequência respectivamente de $x(t)$, $h(t)$ e $y(t)$ (Figura 4.1). Sabendo que um sistema LTI pode ser completamente descrito pela sua resposta impulsional $h(t)$, a relação apresentada pela equação (4.1) é também válida para o domínio das frequências, na medida em que:

$$Y(f) = X(f) \cdot H(f), \quad (4.44)$$

logo:

$$H(f) = \frac{Y(f)}{X(f)}. \quad (4.45)$$

Para um sinal excitador de densidade espectral de potência constante, como é o caso de ruído branco, $X(f)=I$, ou seja, a resposta em frequência do sistema dada por (4.45) resume-se a

$$H(f) = Y(f), \quad (4.46)$$

isto é, à transformada de Fourier da saída $y(n)$ captada pelo microfone.

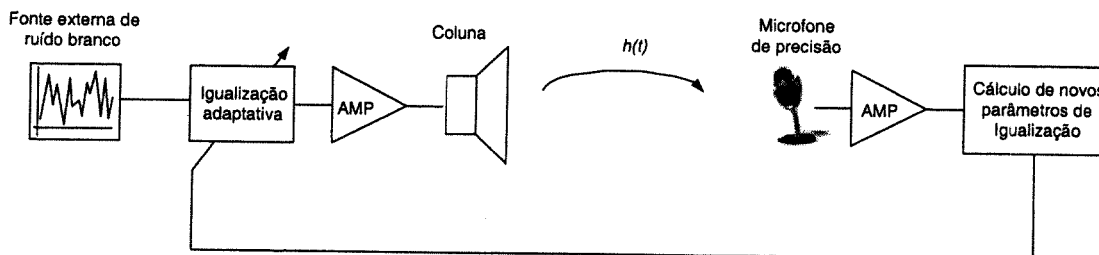


Figura 4.6 - Esquema de excitação por ruído branco

Implementação em DSP

O algoritmo implementado é ilustrado na Figura 4.6, onde se considera que o ruído branco usado na excitação dos modos da sala, é gerado por uma fonte externa ao sistema DSP. Este recebe-o, processa-o e entrega-o ao sistema de amplificação para ser posteriormente aplicado à sala, através de uma coluna de som. Os modos da sala vão ser excitados, sendo o som daí resultante captado pelo microfone de medida e adquirido pela placa de aquisição áudio que, trabalhando sincronamente, o entrega à plataforma DSP. É posteriormente analisado pelo algoritmo que, em cada iteração, vai recalculando adaptativamente novos parâmetros de equalização, a aplicar à entrada do sistema. Estes, vão por sua vez alterar a equalização da sala, tendo como objectivo final uma distribuição 'flat' da energia do sinal pela gama de frequências em análise – a gama audível pelo ser humano, que se situa entre os 20Hz e os 20kHz.

Os coeficientes do filtro FIR são recalculados em cada iteração e mantidos fixos durante o processamento de um bloco inteiro de áudio. Seja n uma indicação temporal do bloco e $w_n(k)$ os coeficientes do filtro para o bloco n . Os coeficientes seguintes $w_{n+1}(k)$ serão obtidos com base nos anteriores da forma descrita em (4.42). G_{est} representa a média espectral a remover de $E\{y(k)\}$ (4.30) no sentido de eliminar a influência do volume na adaptatividade do sistema.

O gradiente $E\{y(k)\} - G_{est}$ dá-nos uma indicação de qual deverá ser a tendência de ajuste dos coeficientes (no sentido inverso ao gradiente, daí o sinal negativo no segundo termo da equação) e da evolução do erro, com base nos resultados obtidos pelo processamento anterior.

O bloco da Figura 4.6 identificado por 'Cálculo de novos parâmetros de equalização' é detalhado na Figura 4.7 para melhor compreensão.

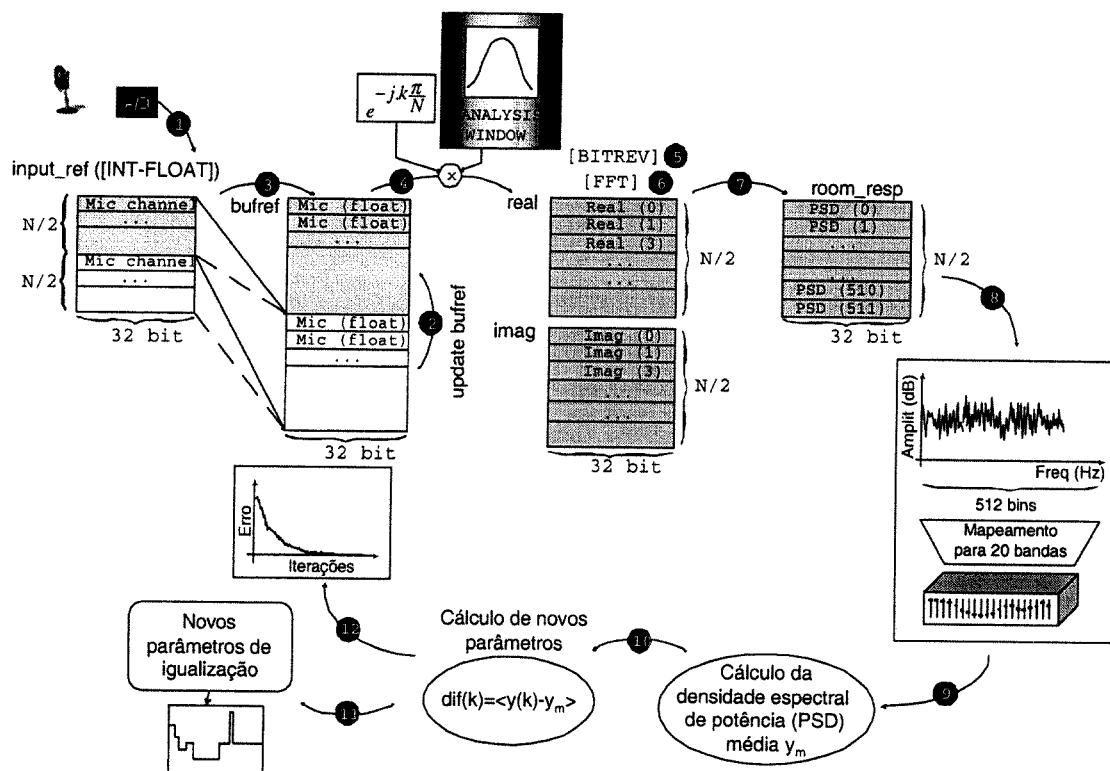


Figura 4.7 - Ajuste adaptativo de novos parâmetros de equalização

Esta parte do algoritmo é responsável pelo cálculo adaptativo de novas curvas de equalização e tem como objectivo minimizar o erro entre a resposta da sala e a resposta desejada 'flat'. Descreve-se de seguida cada um dos passos que a compõem:

1. Baseia-se num mecanismo de captura e armazenamento de áudio utilizando para o efeito um microfone de medida Audix TR40, de elevada qualidade e precisão. As amostras são capturadas pela placa de aquisição áudio AD1836 e enviadas via porta série ao DSP a uma

cadência de 48kHz, onde são armazenadas através de um sistema de *bufferização* circular na memória externa da plataforma DSP – processo descrito pelo passo 1.

2. Metade do buffer '*bufref*' contém amostras do processamento anterior. A sua utilização no processamento actual permitirá a implementação do método de *Overlap-and-Add*. Por esse motivo se actualiza a sua posição no buffer, eliminando a metade mais antiga.
3. A metade deixada livre no buffer é actualizada com as novas amostras recém chegadas da entrada.
4. O conjunto de amostras a processar é multiplicado por uma exponencial, o que resultará no cálculo da ODFT⁴ [Ferreira98]. A posterior multiplicação por uma janela de análise deve-se à utilização do método de *Overlap-and-Add* e ao objectivo de conseguir uma melhor análise espectral.
5. *Bit-reversing* precede o cálculo da transformada rápida de Fourier.
6. A transformada foi totalmente implementada em linguagem Assembly do DSP SHARC 21065L, daí advindo grandes ganhos computacionais relativamente à versão optimizada em linguagem C. É uma transformada complexa, radix-2, com decimação nos tempos (DIT). O processamento é realizado *in-place* uma vez que o DSP não possui tanta memória interna como seria desejável.
7. Com base na informação espectral obtida através da FFT, calcula-se a densidade espectral de potência do sinal.
8. A densidade espectral encontra-se distribuída por 512 bins. Faz-se um mapeamento desses 512 bins em 20 bandas, predefinidas pelo algoritmo de igualização implementado (igualizador digital de 20 bandas).
9. Obtém-se de seguida a densidade espectral média das 20 bandas através da equação
$$\frac{(MAX + MIN)}{2}.$$
10. Primeiro passo na obtenção da nova curva de igualização, com base no estudo dos valores de densidade espectral disponíveis (em dBs) para cada uma das bandas e no seu valor médio. A curva a aplicar terá semelhanças com o inverso da diferença entre estas duas últimas, a menos de um factor de escala.
11. Envio da nova curva à entrada do sistema, que a passará a utilizar a partir desse momento.
12. Cálculo do erro que deverá diminuir com o aumento do número de iterações.

Verifica-se que este método revela uma convergência relativamente rápida do algoritmo para uma curva de igualização estável, que será inversa à da resposta da sala avaliada nas diversas bandas do igualizador gráfico. Como fonte externa geradora de ruído branco, usou-se a aplicação *CoolEdit* para *Windows98* possuidora de um algoritmo próprio gerador de números aleatórios baseado em sementes, e não em algoritmos MLS. Este facto é extremamente

⁴ O cálculo da ODFT poderá ainda ser optimizado, dispensando a multiplicação pela exponencial complexa [Ferreira98].

importante, na medida em que a implementação de um algoritmo deste género será descrita mais tarde e as respectivas comparações serão estabelecidas. A sequência gerada no *CoolEdit* executa em *loop*, e uma vez que é relativamente longa, permite-nos desprezar o efeito da sua periodicidade.

Sendo a fonte geradora de ruído externa ao sistema DSP, torna-se necessário que este o receba. Tal acontece a partir da ligação da placa de som de um PC, à placa de aquisição áudio AD1836 que digitaliza o sinal e o disponibiliza de forma síncrona ao DSP através de uma das suas portas série, para posterior processamento.

Implementou-se um algoritmo de correcção adaptativa, cuja analogia com o método LMS *standard* foi descrita na secção 4.2.3. Realiza-se filtragem adaptativa por blocos no domínio das frequências (FDAF). Os dados de entrada são agrupados em blocos e a filtragem é feita trama a trama, ao invés de o ser amostra a amostra como no caso do algoritmo LMS *standard*.

Critérios de convergência dos parâmetros de igualização

Como se depreende de (4.30), o valor médio da gama dinâmica do sinal $y(k) - G_{est}$, varia em cada iteração. Tal flexibilidade assegura a capacidade do algoritmo alisar o espectro do sinal, actuando de forma mais notória nas bandas mais distantes da média, ou seja, naquelas que mais precisam de correcção. A resposta da sala deverá implicar correcções numa qualquer banda, dentro da gama dinâmica de actuação do igualizador – os $\pm 12\text{dB}$. A Figura 4.8 ilustra uma situação onde se exemplifica a tendência de correcção em cada banda.

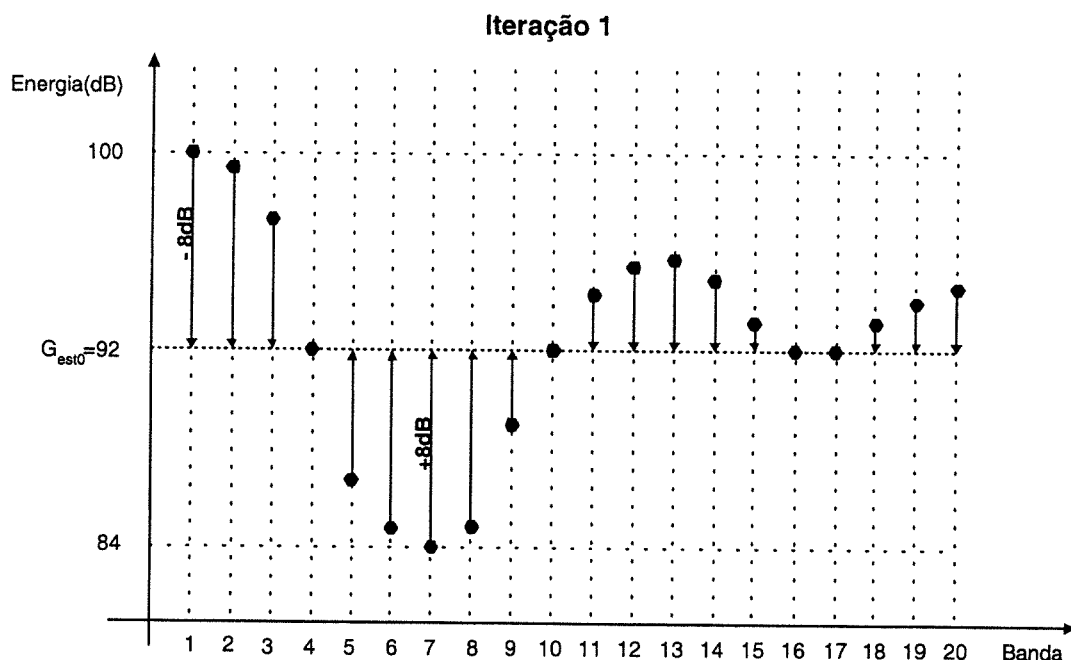


Figura 4.8 – Ajuste em torno de MAX e MIN de $y(k)$

Na situação ilustrada, mostra-se a potência capturada em $y(k)$ para cada uma das 20 bandas. Possui variações entre $+8$ e -8dB , o que garante ao igualizador capacidade de resposta, já que esta se encontrará algures dentro da sua gama dinâmica disponível. Assim, durante a primeira

iteração o algoritmo constata que a potência na banda 1 é máxima (100dB). Do mesmo modo, verifica que a banda 7 é aquela que possui menos potência (84dB) de entre todas as 20 bandas. Como o algoritmo ajusta em função da diferença entre estas duas bandas, verifica-se para a primeira iteração, que o valor médio de ajuste (objectivo) para todas as bandas, deverá ser metade da soma entre esses dois valores de potência (4.30), ou seja, $G_{est0}=92\text{dB}$. Deste modo, a banda 1 e a banda 7 são aquelas que vão sofrer maior ajuste, sendo este respectivamente de -8dB e $+8\text{dB}$. No que respeita às restantes 18 bandas, verifica-se que a cada uma deverá ser aplicado um ganho ou atenuação menores, dependendo de estarem acima ou abaixo do valor médio, para que convirjam em torno de G_{est0} .

Com efeito, no início da segunda iteração provavelmente ainda não haverá um compromisso aceitável entre a resposta '*flat*' desejada e a efectivamente obtida. As duas bandas referidas anteriormente sofreram maior ajuste e poderão não voltar a deter os valores máximo e mínimo de potência em futuras iterações. Assim, urge calcular um novo valor de média espectral, agora em G_{est1} , e a nova adaptação dá-se em torno deste valor para as 20 bandas. Tal situação repetir-se-á iterativamente, até que por ajustes sucessivos na igualização, o espectro convirja para um valor '*flat*' nas 20 bandas em análise, o que de facto acontece.

A vantagem do algoritmo ajustar em torno de um valor médio variável e não de um valor fixo, reside na necessidade de não ser ultrapassada a gama dinâmica do equalizador, ou seja, de não se entrar numa situação em que o ganho ou atenuação a aplicar a uma qualquer banda se situe fora do intervalo $\pm 12\text{dB}$, o que poderia acontecer se o valor médio fosse constante.

De realçar o facto de não se proceder ao cálculo de G_{est} através de uma média aritmética simples da potência nas 20 bandas. Tal motivo deve-se ao possível aparecimento de situações onde há grandes diferenças entre a potência numa das bandas e todas as outras, como se exemplifica na situação limite da figura seguinte:

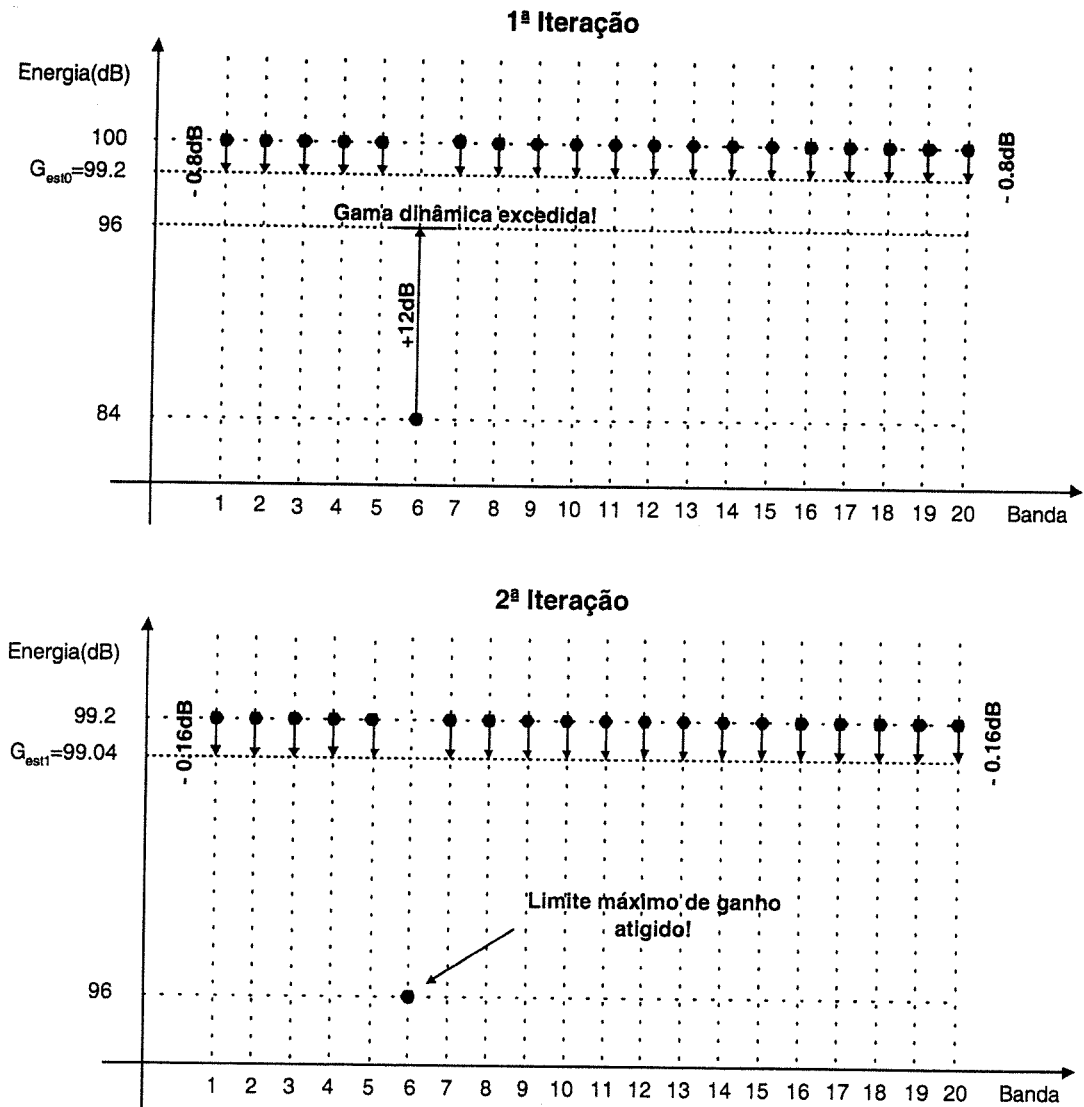


Figura 4.9 - Ajuste em torno da média aritmética de $y(k)$

A banda 6 apresenta um nível de potência muito inferior ao de todas as outras. Com efeito, na primeira iteração, o valor médio G_{est0} seria de 99.2dB, ou seja, $99.2 - 84 = 15.2$ dB acima do valor da banda 6. Nesta situação, a gama dinâmica de actuação do igualizador seria excedida se se usasse o critério de média aritmética no cálculo de G_{est} , e apenas se conseguiria +12dB de ganho a essa banda. Na iteração seguinte, seria calculado um novo valor médio ainda muito distante dos 96dB que agora a banda 6 possui, pelo que seriam apenas as restantes 19 bandas a procurar a convergência do algoritmo, uma vez que a banda 6 não poderia ser corrigida acima dos +12dB. A partir daqui, seriam sempre as outras 19 bandas a procurar com decrementos muito pequenos a aproximação à média G_{est} . Como resultado deste problema, teríamos a convergência do algoritmo num número de iterações elevado.

Ensaio sobre passo de adaptação e médias espectrais

Como referido anteriormente, o sistema gera ruído branco com o qual excita a sala. A resposta em frequência inicial da sala, obtida de acordo com a configuração da Figura 4.6, para uma situação em que ainda não houve qualquer correcção, corresponde à ilustrada na Figura 4.10.

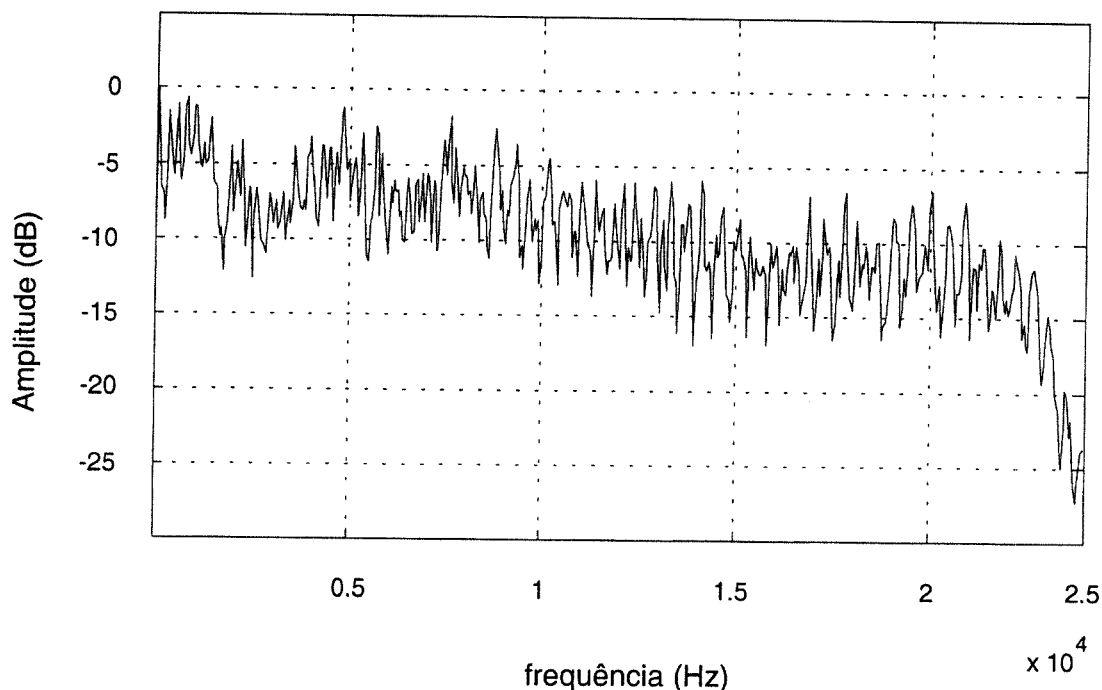


Figura 4.10 - Resposta em frequência inicial da sala

Note-se que a escala de amplitudes não vem referida a valores SPL – *Sound Pressure Level*, que indicam variações de pressão provocadas por uma onda sonora, relativas à pressão atmosférica de $2 \times 10^{-5} \text{ N/m}^2$. É antes uma grandeza logarítmica, normalizada em função do valor máximo da resposta em frequência.

Os dois vales centrados em 2.5 e 6.5 kHz são também visíveis na Figura 4.5 (resposta captada a partir do disparo de uma pistola). Pretende-se que o algoritmo anule estes vales, bem como o vale situado a mais altas frequências (em 16kHz), e que ao mesmo tempo atenuar os picos existentes, de forma a corrigir o espectro para a forma desejada, que se quer 'flat'. Estes diferentes níveis de amplitude verificados devem-se a cancelamentos e reforços introduzidos por reflexões do sinal nas superfícies da própria sala. Como pode ser visto na figura, a potência do sinal atinge entre algumas frequências variações de amplitude superiores a 17dB.

Os testes descritos de seguida, são realizados para diferentes valores da constante μ . Atinge-se um regime estacionário de funcionamento, onde o erro é mínimo, a uma velocidade de convergência que varia com o passo de adaptação. Representa-se na Figura 4.11 a evolução do erro quadrático médio em função do número de iterações, para três valores diferentes de μ .

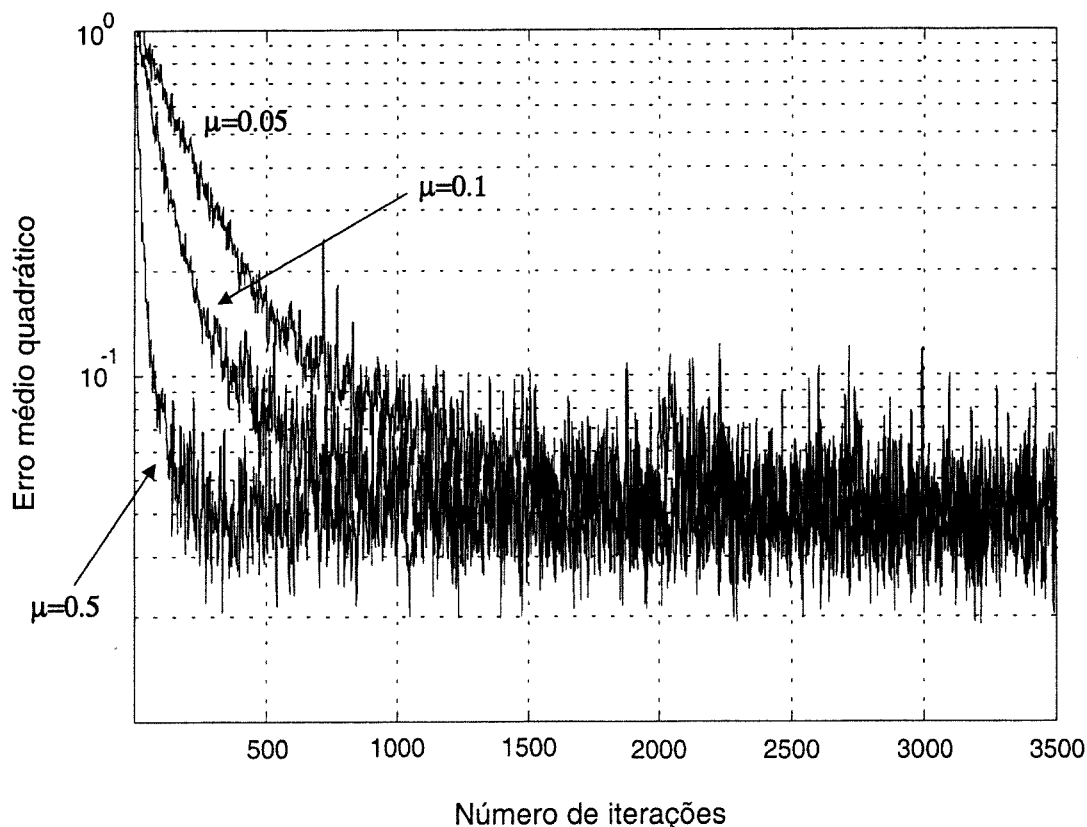


Figura 4.11 - Curva de aprendizagem para o método 4.2.4.2

O erro possui elevada variância, por esse motivo a sua representação gráfica é muito ruidosa. Realizaram-se três experiências para valores de μ , respectivamente 0.05, 0.1 e 0.5. O resultado é o demonstrado na figura e permite concluir que, valores de μ elevados permitem uma convergência mais rápida, como é o caso de $\mu=0.5$. Para $\mu=0.1$, a velocidade de convergência é menor, enquanto que no caso do menor passo, e.g. $\mu=0.05$, verifica-se a mais lenta convergência do três, uma vez que apenas ao fim de cerca de 2000 iterações se obtêm valores de erro próximos dos relativos às experiências anteriores. Por inspecção da Figura 4.11 pode-se observar que o erro converge para valores muito próximos em qualquer um dos três casos, ao fim das primeiras 2000 iterações.

Recapitulando a analogia estabelecida entre o método LMS *standard* e o método implementado e descrito neste capítulo, convém recordar que neste último caso o erro é uma grandeza espectral e não temporal.

O passo de adaptação μ influencia a rapidez de convergência do algoritmo. No método LMS *standard*, um passo μ elevado tem como custo a diminuição da precisão de convergência. Por outras palavras, quanto maior for o passo de adaptação, mais rapidamente o algoritmo convergirá (dentro de certos limites) mas, maior será o erro final, entenda-se, em regime estacionário. Por outro lado, usando aquele método, um passo de adaptação pequeno tem a desvantagem de necessitar de um número superior de iterações para convergir, ao mesmo tempo que apresenta a vantagem de, quando atingido o regime estacionário, o erro de convergência final ser menor. No método implementado, o passo influencia apenas a rapidez

de convergência, uma vez que os níveis de erro final parecem sofrer o efeito de variância do sinal excitador, que se traduz numa variância do sinal captado pelo microfone, não permitindo que o algoritmo convirja para baixo de um determinado erro. Adicionalmente, efeitos físicos de natureza acústica dentro da sala, parecem também exercer alguma influência no aparecimento da referida variância. Por outras palavras, as variâncias provocadas pelo sinal excitador e por influência da sala são superiores à variância que o passo μ introduz.

A correcção introduzida na sala para o melhor dos casos, resulta de (4.42) ter adaptativamente encontrado o melhor conjunto de níveis de igualização, ilustrado na figura de baixo:

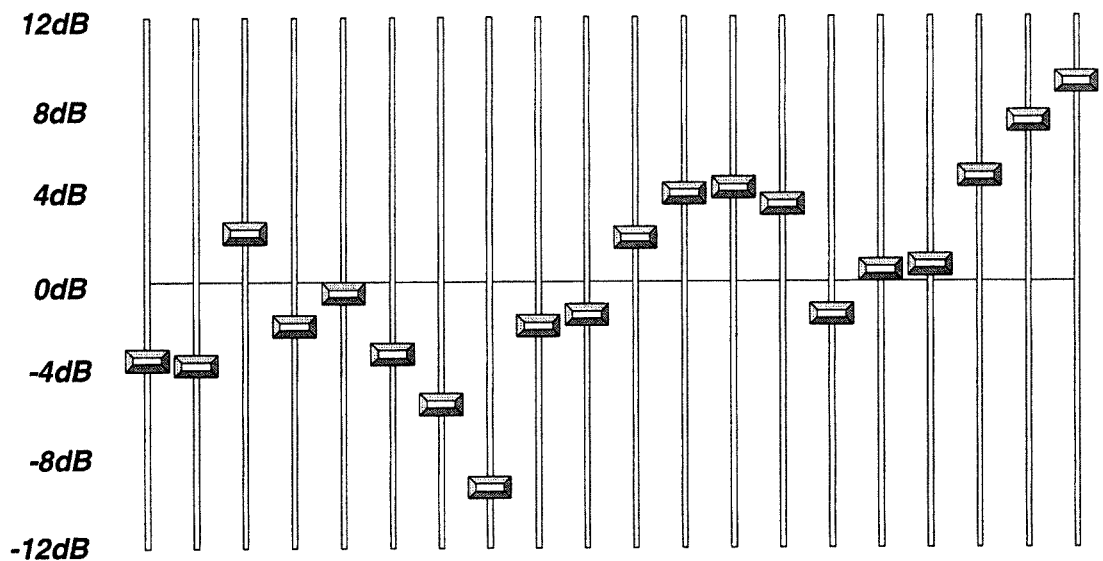


Figura 4.12 – Níveis de igualização da sala obtidos adaptativamente

A Tabela 1 indica aqueles valores com detalhe, acompanhados da correspondente frequência central de cada banda de controlo do equalizador.

Variando o volume do sistema de amplificação, verifica-se que os níveis de equalização mantêm o seu perfil, com uma variação pontual não muito significativa numa das bandas, fruto do efeito do aumento da energia de excitação num dos modos de ressonância da sala, cujo comportamento não se mostra tão linear como seria à partida de esperar.

BANDA	CENTRO (Hz)	NÍVEIS DE IGUALIZAÇÃO
1	47	-3.4
2	140	-3.6
3	234	2.4
4	352	-2
5	492	-0.6
6	633	-3.4
7	820	-5.4
8	1008	-9.3
9	1242	-2
10	1523	-1.5
11	1875	2.1
12	2320	4.1
13	2930	4
14	3680	3.5
15	4805	-1.6
16	6352	0.6
17	8507	0.6
18	11414	4.6
19	15210	7.2
20	20883	9.1

Tabela 1 - Níveis de igualização final

O espectro final corrigido pelo sistema adaptativo, tornou-se como pretendido desde o início, 'flat'. Compare-se o seu aspecto representado na Figura 4.13, com o da Figura 4.10. Os vales foram suprimidos, bem como as zonas de maior amplitude, ao mesmo tempo que a potência do sinal sofreu um ajuste em torno do valor médio ideal, que parece neste caso rondar os -7dB (valor relativo). A oscilação dos níveis de potência visível na Figura 4.10 e Figura 4.13, deve-se à existência de ecos e à forte variância do sinal excitador, que é ruído branco. Efectuando a média a um número superior de tramas áudio, seria possível alisá-lo um pouco mais, apesar de tal alteração provocar um aumento exponencial no tempo de processamento necessário para convergência do algoritmo (note-se pela Figura 4.11 que se realizaram mais de 4500 iterações para cada teste).

Os testes foram realizados efectuando a média dos espectros de 50 tramas do sinal áudio captado pelo microfone, para posterior cálculo adaptativo dos coeficientes, como sugere o valor esperado $E\{y(k)\}$ da equação (4.42). A realização de médias justifica-se pela variância do espectro do sinal excitador (ruído branco), que se apresenta muito recortado e irregular, por vezes com vales profundos. Variando o número de médias dos espectros capturados pelo microfone, verificam-se alterações no erro em regime estacionário. Com efeito, realizando

apenas 10 médias espectrais aumenta-se a velocidade de convergência, à custa de um aumento significativo do erro (o dobro) em regime estacionário. Se, por outro lado, se realizar a média de 200 espectros, observa-se que o erro diminui bastante (para cerca de metade), mas apresenta a contrapartida de o algoritmo convergir num intervalo de tempo mais elevado, ou seja, diminui-se a velocidade de convergência. Nesta situação, os níveis e igualização encontrados aproximam-se com maior precisão dos ideais, pelo que os ajustes finais são já ajustes finos. Nenhuma das alterações anteriores provocou, no entanto, uma alteração significativa no perfil de igualização, o que é desejável. A solução escolhida de realizar a média de 50 segmentos espectrais apresenta, assim, um bom compromisso entre erro final que se deseja pequeno (ajuste fino), e velocidade de convergência que se pretende elevada (níveis de igualização ideal da sala obtidos em poucos segundos).

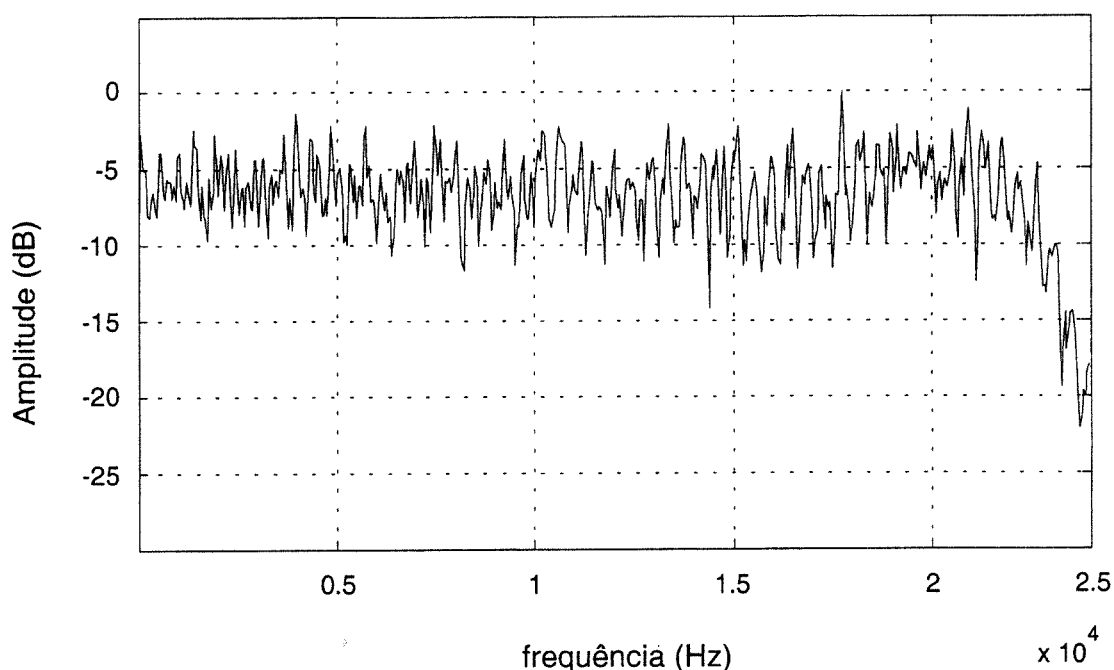


Figura 4.13 - Resposta em frequência final da sala

Consegue-se desta forma a correcção do espectro, que no caso do passo de adaptação $\mu=0.5$, se pode considerar ter convergido ao fim de 100 iterações, o que para o número de média de tramas calculado equivale a cerca de 50 segundos. A partir daí, os ajustes efectuados são menos significativos. No que diz respeito ao passo menor, i.e. $\mu=0.05$, verifica-se que são necessárias cerca de 2000 iterações, i.e. cerca de 17 minutos, para atingir um valor de erro de convergência que não é significativamente inferior ao obtido pelo passo anterior. A primeira solução parece portanto ser a mais eficaz. Do ponto de vista prático, há o interesse em reduzir o tempo de convergência para um valor consideravelmente inferior a 50 segundos. Tal tarefa poderá ser conseguida aceitando um erro final (em regime de convergência) superior, desde que aumento não seja significativo, ou excitando a sala com um tipo de técnica que não exija o cálculo de tantas médias espectrais, na obtenção da resposta em frequência da sala.

O gráfico da Figura 4.14 representa um mapeamento da Figura 4.13 para as 20 bandas que o sistema consegue controlar.

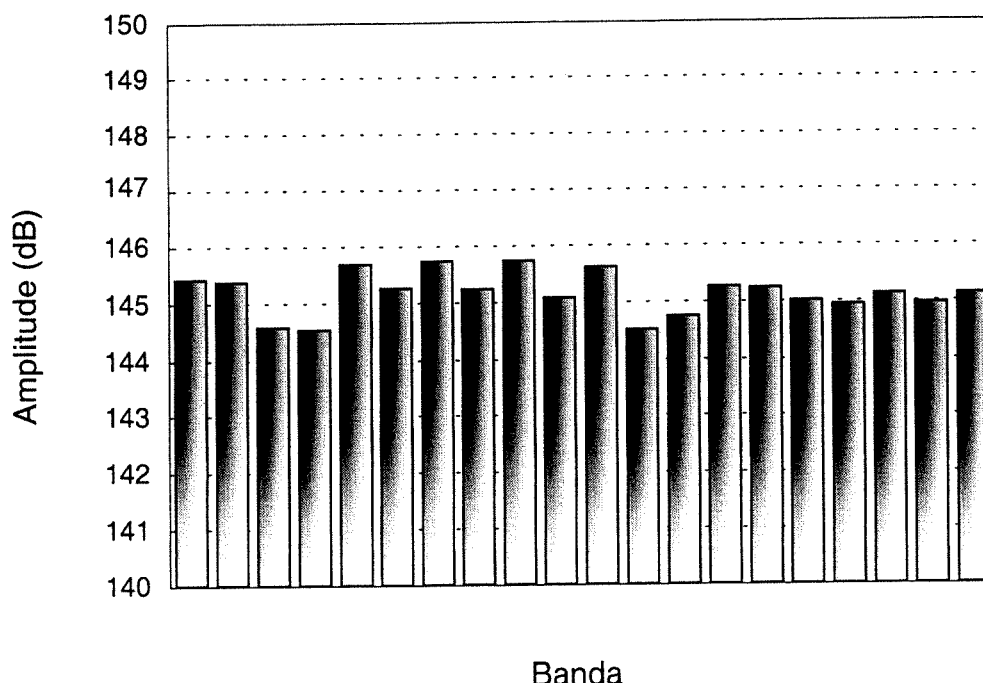


Figura 4.14 - Mapeamento da resposta em frequência nas 20 bandas

Novamente, a escala de amplitudes no eixo das ordenadas não vem referida a valores SPL – *Sound Pressure Level*. É antes uma grandeza logarítmica, afectada pelos ganhos introduzidos no sinal no processo de captura da resposta da sala.

O variação em torno do valor médio ideal não é nula, o que implica que o alisamento espectral não seja perfeito. Verificam-se fenómenos de inconstância na amplitude do sinal que excita a sala, que causam a variação do espectro (visível na Figura 4.14) em torno de um valor médio ideal. Esta variação não está relacionada com problemas de adaptatividade do algoritmo em si, mas antes com fenómenos acústicos da sala (e.g. variações nas condições de eco) e com a variância do sinal excitador. Variações temporais poderão surgir nas medidas efectuadas em várias situações: na realização de testes ao ar livre, onde o som percorre grandes distâncias; quando são realizadas médias do som capturado ao longo de um período de tempo considerável; quando o sistema não é suficientemente invariante nos tempos. Em situações onde é por vezes necessário efectuar centenas ou até milhares de médias sobre o sinal capturado (que envolvem longos períodos de captura), uma ligeira variação de temperatura ou movimento do ar pode introduzir perturbações no cálculo da média [Svensson&Nielsen99]. Por outro lado, é conhecido que sequências de ruído periódicas são extremamente vulneráveis a variações temporais [Müller&Massarini2001]. Na prática, é difícil extrair conclusões acerca dos efeitos que provocam variação temporal, na medida em que esta tende a possuir uma natureza errática e de carácter imprevisível. Imagine-se a título exemplificativo, duas medidas efectuadas ao ar livre (por exemplo num estádio) de forma consecutiva, em que os dois

resultados obtidos vêm afectados de forma diferente por variações na influência provocada pelo vento.

4.3.3. ALGORITMO COM EXCITAÇÃO ATRAVÉS DE SEQUÊNCIAS MLS.

Os sinais MLS (*Maximum Length Sequences*) são sequências binárias pseudo-aleatórias sintetizadas digitalmente, possuidoras de propriedades matemáticas interessantes, que se pretendem explorar no âmbito do presente trabalho.

As sequências MLS possuem várias qualidades: têm um factor de *crest*⁵ mínimo, o que significa que transferem muita energia para o sistema e por esse motivo obtêm-se medidas com óptima SNR. Como o estímulo é determinístico, proporciona repetitividade exacta. É um método apropriado para a medida de respostas impulsivas longas.

Com o uso de análise MLS procura-se caracterizar o comportamento acústico de uma sala, no domínio das frequências, através do estudo da transformada da sua resposta impulsional (IR), com o objectivo de lhe aplicar um pós-processamento para correcção de perturbações, nomeadamente efeitos provocados por reflexões do sinal nas superfícies da sala.

É possível alimentar a entrada de um sistema com uma sequência MLS, para através do cálculo da sua correlação cruzada com a resposta deste, obter a resposta impulsional (IR) do sistema. Uma forma eficiente de realizar o cálculo de correlação cruzada entre dois sinais, consiste em utilizar a Transformada Rápida de Hadamard (FHT), onde uma correlação cruzada de segmentos de comprimento L é calculada recorrendo apenas a $2.5L \log_2 L$ adições em vírgula flutuante.

Com o aparecimento de DSPs rápidos e de baixo custo e de conversores digitais-analógicos de elevada qualidade e igualmente baixo custo, tornou-se mais eficiente gerar este tipo de sinais via *software*, o que se consegue recorrendo a menos de 20 linhas de código.

Algoritmo gerador de sequências MLS

Descreve-se nesta secção o processo de geração de sequências MLS e abordam-se as propriedades matemáticas que as caracterizam. Existem sequências MLS de estados ternários ou superiores, se bem que a implementação realizada no decorrer deste trabalho e aqui descrita, se limita a sequências binárias.

O sistema gerador de sequências binárias é composto por um *feedback shift register* de N andares (N elementos binários) e por uma *gate XOR* realimentada para a entrada, onde são gerados $2^N - 1$ estados, i.e., todos os estados possíveis menos o estado 'tudo zeros' (Figura 4.15). Da escolha dos elementos binários que o *shift register* fornece como entrada à *gate XOR*, depende o facto da sequência gerada ser de comprimento máximo (como pretendido), ou não. Dependendo do número de andares do *shift register*, é possível escolher um polinómio

⁵ É definido como sendo a razão entre o valor de pico e o valor eficaz (RMS) de um sinal. Um factor de *Crest* elevado indica que o sinal possui desvios raros, mas muito acentuados, relativamente aos seus níveis normais.

primitivo, que indique os elementos binários a utilizar, de modo a que a sequência gerada possua comprimento máximo, i.e., todos os 2^N-1 possíveis estados sejam gerados, antes que o sinal se repita. Para um mesmo registo de dimensão N , é em muitos casos possível escolher mais do que um polinómio primitivo, sendo que cada um desses polinómios gerará igualmente uma sequência de comprimento máximo, só que completamente descorrelacionada com a gerada pelo polinómio anterior, ou seja, para o mesmo número N de andares podem gerar-se diferentes sequências com o mesmo comprimento.

Na realização deste trabalho, implementou-se um *shift register* de 17 andares, como ilustrado na Figura 4.15, sendo $P(x)=x^{17}+x^{12}+1$ o polinómio escolhido [Peterson&Ziemer&Borth95], que garante a geração de uma sequência de comprimento máximo. Com $N=17$ andares geram-se $2^{17}-1=131071$ estados não nulos, que à frequência de amostragem 48kHz originam uma sequência com 2.7seg de duração.

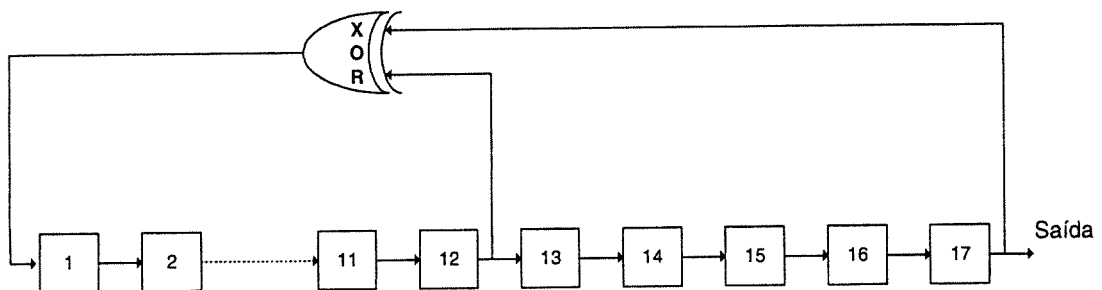


Figura 4.15 - *Shift register* gerador de sequências MLS

Os elementos indicados pelo polinómio primitivo são somados através de uma função *Or-Exclusivo* (XOR) e realimentados para o primeiro elemento do *shift register*, depois de deslocar todo o conteúdo do mesmo um elemento em direcção à saída. A saída é produzida no último elemento do *shift register*. O ruído gerado aproxima-se de certo modo a ruído branco, mas na realidade é determinístico (conhecido em cada instante) e periódico, com um comprimento que depende da ordem N do *shift register*. Por esse motivo é denominado ruído pseudo-aleatório. Um registo de N andares é capaz de gerar 2^N estados. O *shift register* implementado é capaz de operar em 2 modos: gerando o estado nulo onde todos os elementos do registo contêm zeros; ou gerando todo o outro conjunto de 2^N-1 estados. No primeiro caso, o algoritmo evolui sempre para estados '0', uma vez que $0 \text{ XOR } 0=0$. No segundo, percorre todos os restantes 2^N-1 estados não nulos. Nenhum estado não nulo é capaz de gerar o estado nulo. Por esse motivo, o registo deve ser inicializado num qualquer valor diferente de zero. É típico inicializá-lo no estado 'tudo 1s'.

Concatenando os bits de saída do *shift register*, obtém-se uma sequência de comprimento 2^N-1 . O número de '1s' numa sequência de comprimento máximo excede o número de '0s' exactamente em 1 unidade, o que implica que a soma binária de uma sequência deste género resulte em 2^{N-1} .

Em aplicações práticas, a saída binária do registo '0' e '1' é normalmente mapeada, respectivamente, para os níveis eléctricos simétricos '+1' e '-1', o que permite designar o sinal por MLS simétrico. Pela característica relativa à quantidade de '0s' e '1s' presentes numa sequência deste género, descrita anteriormente, é possível verificar que a soma dos elementos

de um sinal MLS simétrico resulta em ‘-1’, o que permite concluir que este tipo de sinais contém um nível de energia DC muito reduzido.

Propriedades da Autocorrelação e Correlação Cruzada

Recorde-se uma vez mais a Figura 4.1. Para sinais contínuos a função de autocorrelação $r_{xx}(t)$ de uma determinada entrada $x(t)$ é dada por:

$$r_{xx}(t) = \int_{\tau=0}^{\infty} x(\tau)x(\tau+t)\partial\tau \quad (4.47)$$

enquanto que a correlação cruzada entre a entrada $x(t)$ e a saída $y(t)$ do sistema, é dada por:

$$r_{xy}(t) = \int_{\tau=0}^{\infty} x(\tau)y(\tau+t)\partial\tau. \quad (4.48)$$

O integral de convolução:

$$y(t) = \int_{\lambda=0}^{\infty} x(\lambda)h(t-\lambda)\partial\lambda \quad (4.49)$$

define a saída $y(t)$ do sistema em função da sua resposta $h(t)$ e da sua entrada $x(t)$, e pode ser utilizado em (4.48) para relacionar a função de autocorrelação com a correlação cruzada, originando:

$$r_{xy}(t) = \int_{\tau=0}^{\infty} x(\tau) \int_{\lambda=0}^{\infty} x(\lambda)h(\tau+t-\lambda) \partial\lambda \partial\tau. \quad (4.50)$$

Aplicando a mudança de variável $\lambda = \tau + \gamma$ a equação resulta em:

$$r_{xy}(t) = \int_{\gamma=0}^{\infty} \int_{\tau=0}^{\infty} x(\tau)x(\tau+\gamma)h(t-\gamma) \partial\tau \partial\gamma \quad (4.51)$$

Analisando (4.47), obtém-se (4.51) escrita da forma:

$$r_{xy}(t) = \int_{\gamma=0}^{\infty} r_{xx}(\gamma)h(t-\gamma)\partial\gamma, \quad (4.52)$$

ou,

$$r_{xy}(t) = r_{xx}(t) * h(t) \quad (4.53)$$

Sinais de potência espectral constante, como é o caso de ruído branco, possuem autocorrelação semelhante a um impulso de *Dirac*. Assim sendo, (4.53) pode ser representada por:

$$r_{xy}(t) = \delta(t) * h(t), \quad (4.54)$$

o que não é mais do que:

$$r_{xy}(t) = h(t), \quad (4.55)$$

o que prova que para sinais excitadores $x(t)$ deste género (de potência espectral constante), a resposta impulsional $h(t)$ pode ser directamente obtida pelo cálculo da correlação cruzada entre a entrada $x(t)$ e a saída $y(t)$ do sistema.

Sistemas periódicos discretos nos tempos

Reportemo-nos uma vez mais à Figura 4.1. Considerando agora que se trabalha com sinais discretos nos tempos, recorde-se que a sequência de saída $y(n)$ de um sistema LTI deste tipo, depende da sequência de entrada do sistema $x(n)$ e da sua resposta impulsional $h(n)$ da forma descrita por (4.1) e (4.2). Usando um impulso unitário $\delta(n)$ como sequência de entrada, onde:

$$\delta(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}, \quad (4.56)$$

a resposta do sistema $y(n)$ torna-se igual à resposta impulsional unitária $h(n)$, que descreve por completo o sistema.

Se em alternativa ao impulso unitário anterior se usar agora uma sequência de entrada $\delta'(n)$ semelhante, mas periódica [Rife&Vanderkooy89]:

$$\delta'(n) = \begin{cases} 1, & n \bmod L = 0 \\ 0, & n \bmod L \neq 0 \end{cases}, \quad (4.57)$$

a relação entre a resposta do sistema $y'(n)$ em função da sua resposta impulsional $h'(n)$ e da sequência de entrada $x'(n)$ altera-se, na medida em que passa a ser descrita pela convolução periódica:

$$y'(n) = x'(n) * h'(n), \quad (4.58)$$

ou,

$$y'(n) = \sum_{k=0}^{L-1} x'(k) h'(n-k), \quad (4.59)$$

onde todos os sinais são grandezas periódicas, de período L . A relação entre a resposta impulsional anterior $h(n)$ e a resposta impulsional periódica $h'(n)$ pode ser descrita por:

$$h'(n) = \sum_{k=-\infty}^{+\infty} \delta'(k)h(n-k). \quad (4.60)$$

De facto, $\delta'(n)$ tem o efeito de envolver $h(n)$ em torno de si própria, acumulando segmentos de L pontos, originando a resposta impulsional periódica $h'(n)$:

$$h'(n) = \sum_{k=-\infty}^{+\infty} h(n+kL). \quad (4.61)$$

Para contornar este problema, escolhe-se um período L que será suficientemente grande, se a resposta $h(n)$ decair para um valor desprezável após a sua L -ésima amostra. Minimiza-se deste modo o fenómeno de sobreposição descrito, designado por '*aliasing*' temporal e $h'(n)$ aproxima-se com bastante fidelidade das primeiras L amostras de $h(n)$.

Deve ser tido em consideração que métodos de excitação periódica para medida da resposta impulsional de um sistema, resultam sempre numa resposta impulsional periódica, pelo que L deverá ser escolhido com a devida precaução. Para uma correcta realização desta tarefa, o tempo de decaimento do sistema deverá ser conhecido, ou em alternativa deverão usar-se sequências longas.

Propriedades discretas de sequências MLS

Sequências MLS possuem propriedades matemáticas, cuja manipulação resulta em benefícios que se pretendem aproveitar, na implementação de um sistema de medida da resposta impulsional de uma sala.

De facto, a autocorrelação $r'_{xx}(n)$ de uma sequência MLS simétrica periódica $x'(n)$ continua a assemelhar-se a um impulso unitário, só que agora periódico. Ou seja:

$$r'_{xx}(n) = \frac{1}{L} \sum_{k=0}^{L-1} x'(k)x'(n+k) \quad (4.62)$$

que resulta em [Rife&Vanderkooy89]:

$$r'_{xx}(n) = \begin{cases} 1, & n=0 \\ -\frac{1}{L}, & 0 < n < L \end{cases} \quad (4.63)$$

Usando (4.57), conclui-se [Rife&Vanderkooy89] que (4.63) pode ser representada por:

$$r'_{xx}(n) = \delta'(n) - \frac{1}{L}. \quad (4.64)$$

Verifica-se deste modo que a função de autocorrelação é constituída por um impulso unitário e por um valor DC, que será tanto menor, quanto maior for o período L da sequência [Rife&Vanderkooy89]. Nesta situação, $r'_{xx}(n)$ aproxima-se bastante do impulso unitário ideal

$\delta(n)$. Adicionalmente, a autocorrelação fornece também uma medida do grau de aleatoriedade do sinal.

Uma vez mais, se $x'(n)$ for uma sequência MLS aplicada à entrada de um sistema LTI, a saída $y'(n)$ será a convolução de $x'(n)$ com a resposta impulsional desse sistema $h'(n)$, como indicado em (4.59). Analogamente a (4.55), sabendo que se usa um sinal (MLS) de espectro constante como entrada do sistema, também aqui a resposta impulsional (agora periódica) do sistema pode ser obtida por cálculo da correlação cruzada periódica entre entrada e saída, como se demonstra de seguida:

$$r'_{xy}(n) = \frac{1}{L} \sum_{k=0}^{L-1} x'(k) y'(n+k). \quad (4.65)$$

A parcela $y'(n+k)$ pode ser expandida:

$$y'(n+k) = \sum_{\ell=0}^{L-1} x'(n+k-\ell) h'(\ell), \quad (4.66)$$

o que origina (4.65) da forma:

$$\begin{aligned} r'_{xy}(n) &= \frac{1}{L} \sum_{k=0}^{L-1} x'(k) \sum_{\ell=0}^{L-1} x'(n+k-\ell) h'(\ell) \\ &= \frac{1}{L} \sum_{\ell=0}^{L-1} \sum_{k=0}^{L-1} x'(k) x'(n+k-\ell) h'(\ell) \end{aligned} \quad (4.67)$$

que, por (4.62) se resume a:

$$r'_{xy}(n) = \sum_{\ell=0}^{L-1} r'_{xx}(n-\ell) h'(\ell). \quad (4.68)$$

Concluindo, a correlação cruzada periódica é, como se desejava demonstrar:

$$r'_{xy}(n) = r'_{xx}(n) * h'(n). \quad (4.69)$$

Encontra-se igualmente desta forma, para o caso em que a excitação de entrada é periódica, a relação entre autocorrelação periódica da entrada $x'(n)$ e correlação cruzada periódica entre entrada $x'(n)$ e saída $y'(n)$, o que nos permite obter a resposta impulsional (IR). Recorrendo a (4.64), é possível desenvolver a equação (4.69) no sentido de eliminar parcelas desprezáveis [Rife&Vanderkooy89]:

$$r'_{xy}(n) = \left[\delta'(n) - \frac{1}{L} \right] * h'(n). \quad (4.70)$$

Desenvolvendo mais,

$$r'_{xy}(n) = h'(n) - \frac{1}{L} \sum_{k=0}^{L-1} h'(k), \quad (4.71)$$

observa-se que a parcela do segundo termo da equação dependente do comprimento L da sequência, que representa o valor médio da resposta impulsional periódica, se torna desprezável se a sequência possuir um comprimento L elevado.

Implementação em DSP da técnica MLS

Este segundo método implementado excita sonoramente a sala através de sequências MLS, tirando proveito das características matemáticas anteriormente referidas. Note-se que a resposta de um sistema acústico entre uma fonte sonora e um receptor descreve apenas a resposta entre esses dois locais e não a resposta completa da sala. A sequência MLS é gerada digitalmente pelo sistema DSP, logo é discreta nos tempos. Observa-se na Figura 4.16 o efeito que a mesma sofre, a propósito da sua conversão para o domínio contínuo.

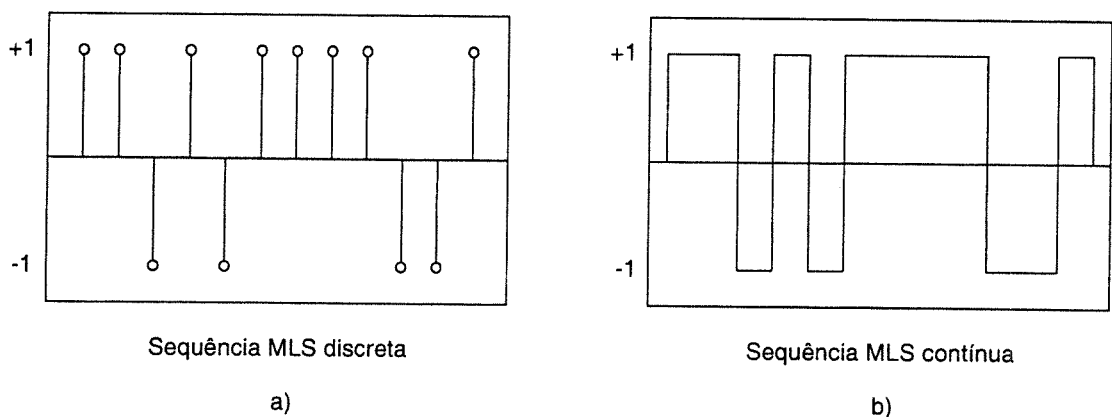


Figura 4.16 - Conversão de uma sequência MLS para o domínio contínuo

A passagem de um domínio para o outro equivale, do ponto de vista de processamento de sinal, à convolução entre impulsos unitários e um pedestal. Ou, fazendo a analogia com o domínio das frequências, equivale a multiplicar o perfil *'flat'* que o espectro da sequência MLS possui, pela função $\text{sinc}(f)^\dagger$, que é como sabido, a transformada do referido pedestal. Esta coloração espectral indesejada da sequência MLS é, no entanto, evitada por conversores digitais-analógicos modernos, que efectuem a compensação do efeito $\text{sinc}(f)$.

O algoritmo implementado gera como foi já referido, uma sequência MLS, fonte de excitação sonora do sistema. O esquema de excitação usado neste método, pode de certa forma ser representado pela Figura 4.6, tendo em consideração que agora a fonte excitadora não é externa ao sistema DSP, uma vez que a sequência é gerada por este, e que no bloco designado por *'Cálculo de novos parâmetros de igualização'* reside a principal diferença entre este método e o anteriormente implementado em 4.3.2. O algoritmo gerador da sequência MLS é, como referido anteriormente, implementado recorrendo a poucas linhas de código:

[†] Definida por $\text{sinc}(f) = \frac{\sin(f)}{f}$.

```

#define XOR ^

for(i=17;i>0;i--)
    MLS_seq[i] = MLS_seq[i-1];

MLS_seq[0] = MLS_seq[12] XOR MLS_seq[17];

if(MLS_seq[17] == 1)
{
    MLS_bit = 1;
    MLS_output = -1;
}
else
{
    MLS_bit = 0;
    MLS_output = 1;
}

```

Figura 4.17 - Algoritmo MLS em linguagem C

É implementado na própria rotina de atendimento a interrupções (ISR) e o valor de saída multiplicado por uma constante, para posterior envio da amostra para a porta série, conversor digital-analógico, sistema de amplificação e coluna de som.

O bloco '*Cálculo de novos parâmetros de igualização*' é responsável pelo cálculo adaptativo dos novos níveis (parâmetros) de igualização a afectar o igualizador, que vão iterativamente corrigindo perturbações acústicas da sala provocadas por reflexões do sinal nas suas superfícies, com o objectivo de tornar o seu espectro '*flat*' para um determinado local daquela, mediante a minimização de uma função de erro entre espectro desejado e espectro obtido. A ilustração do método agora implementado (excitação por MLS) é apresentada pelo diagrama de blocos da Figura 4.18.

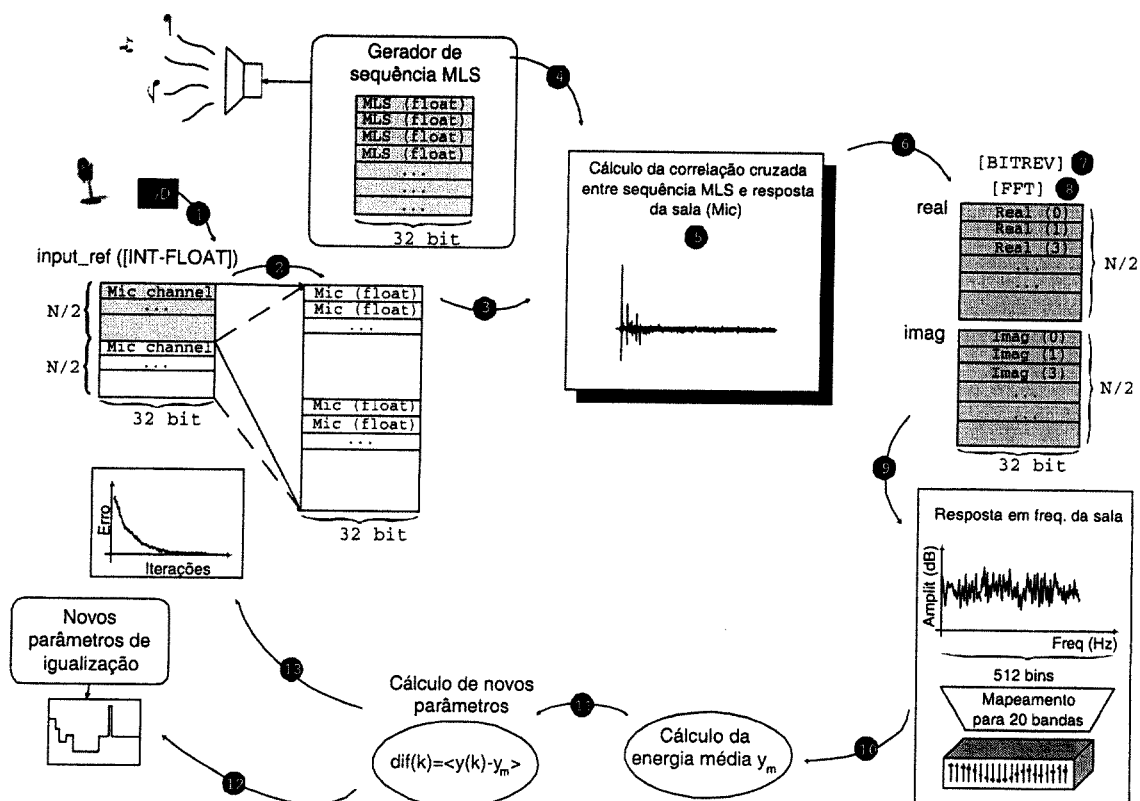


Figura 4.18 - Ajuste adaptativo de novos parâmetros de equalização com excitação MLS

Os modos da sala são excitados por uma sequência de ruído pseudo-aleatório (MLS), e aquela reage de acordo com a resposta dos seus modos ressonantes, num sinal que é captado por um microfone de precisão. Uma descrição detalhada de todo o processamento que envolve este algoritmo é dada de seguida para cada um dos passos representados a negro na Figura 4.18:

1. Mecanismo de captura da resposta áudio da sala através de um microfone de medida *Audix TR40*, de elevada precisão, idêntico ao descrito no método anterior.
2. Armazenamento de amostras vindas da entrada, fruto da captura da resposta da sala pelo microfone, no *buffer* 'micL_buf', que fornece assim um dos sinais necessários para o cálculo da correlação cruzada.
3. Leitura de amostras de entrada (respeitando o devido atraso) para cálculo de correlação cruzada.
4. Geração do sinal MLS e envio do mesmo para o conversor D/A, que o fornece ao sistema de amplificação para com ele excitar a sala. Leitura dessa sequência MLS para cálculo de correlação cruzada.
5. Cálculo de correlação cruzada em rotina implementada em linguagem Assembly, para execução rápida.
6. Deslocamento da resposta impulsional da sala (resultado do cálculo da correlação cruzada) para memória interna do DSP, para cálculo rápido da resposta em frequência da sala, usando Transformadas de Fourier Rápidas (FFTs).

7. *Bit-reversing* precede o cálculo da transformada rápida de Fourier.
8. Do mesmo modo, que no processamento anterior, a transformada foi totalmente implementada em linguagem Assembly do *DSP SHARC 21065L*, o que permitiu obter elevados ganhos do ponto de vista de rapidez de execução, relativamente à versão otimizada em linguagem C. É de novo uma transformada complexa, radix-2, com decimação nos tempos (DIT) e processamento realizado *in-place*, em memória interna (para cálculo rápido).
9. A densidade espectral de potência do sinal representando a resposta em frequência da sala encontra-se distribuída por 512 *bins*. Faz-se neste passo um mapeamento desses valores para as 20 bandas de controlo do algoritmo igualizador.
10. Obtenção da potência média das 20 bandas, utilizando a fórmula $\frac{(MAX + MIN)}{2}$.
11. Através do estudo dos valores de densidade espectral obtidos (em dBs) para cada uma das bandas e do seu valor médio, obtém-se a nova curva de igualização a aplicar ao sistema. A nova curva deverá possuir uma tendência inversa à do espectro actual da sala, no sentido de o corrigir.
12. Aplicação da nova curva à entrada do sistema que passará a excitar a sala com ruído MLS igualizado pela nova curva.
13. Cálculo do erro, usado para monitorizar a convergência do algoritmo. Deverá diminuir com o aumento do número de iterações.

O método de correcção adaptativa implementado possui uma analogia com o método LMS *standard*, descrita no capítulo 4.2.3. Realiza-se filtragem adaptativa por blocos (trama a trama, e não amostra a amostra) no domínio das frequências (FDAF), ficando os coeficientes do filtro fixos durante o processamento de uma trama de áudio completa. De novo, seja n uma indicação temporal e $w_n(k)$ os coeficientes do filtro para o n -ésimo processamento. Os próximos coeficientes $w_{n+1}(k)$ serão calculados com base nos anteriores. No método descrito no capítulo 4.3.2 (excitação por ruído branco), os coeficientes são calculados com base em (4.42), onde a indicação do valor esperado $E\{y(k)\}$ para a resposta da sala capturada pelo microfone, sugere a necessidade de efectuar uma média a vários espectros, antes do cálculo dos novos coeficientes para uma determinada iteração. A necessidade dessa média é justificada pela natureza do sinal excitador (de espectro branco), que se verifica possuir um espectro bastante recortado, por vezes afectado por vales acentuados. Neste segundo método (excitação por técnica MLS), a resposta impulsional da sala é obtida por cálculo da correlação cruzada entre sinal excitador e resposta captada pelo microfone (4.65) e (4.71). Como estes dois sinais possuem 2.7seg de duração (L elevado em (4.71)), assume-se que a resposta impulsional obtida através da correlação cruzada daqueles dois sinais possui duração suficiente, finda a qual, as suas variações de potência possam ser consideradas desprezáveis. A Figura 4.19 revela que apenas para os primeiros 40ms da resposta impulsional da sala, a potência decai de forma considerável.

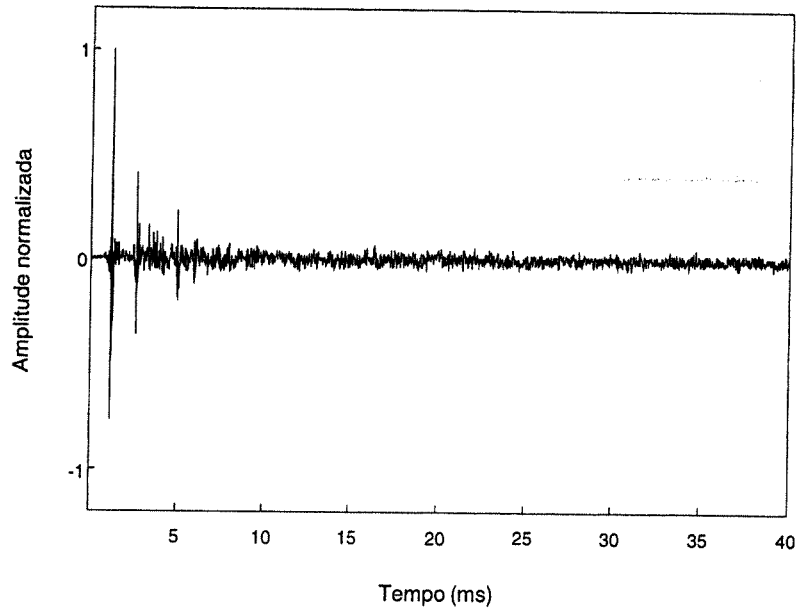


Figura 4.19 - Resposta impulsional da sala para os primeiros 40ms

Assim, como a resposta em frequência da sala é obtida através do cálculo da transformada de Fourier da sua resposta impulsional, para um segmento de tempo considerável ao fim do qual se assume que as variações de potência não introduzirão variações significativas na resposta da sala, conclui-se que não há neste caso a necessidade de efectuar médias espectrais no sinal capturado pelo microfone. Deste modo, a equação usada no cálculo adaptativo dos novos coeficientes é agora dada por (4.43), que aqui se recapitula:

$$w_{n+1}(k) = w_n(k) - \mu[y(k) - G_{est}], \quad (4.72)$$

e cuja diferença relativamente a (4.42) consiste na remoção do valor esperado $E[y(k)]$.

Da mesma forma que no método descrito anteriormente em 4.3.2, também aqui G_{est} é calculado por (4.30) e representa a média espectral (da potência nas 20 bandas) a remover de $y(k)$ no sentido de eliminar a influência do volume na adaptatividade do sistema. Do mesmo modo, G_{est} varia em cada iteração, de forma a permitir ao algoritmo 'alisar' o espectro, actuando de forma mais acentuada nas bandas mais distantes da média desejada⁶.

Testaram-se dois valores para o passo de adaptação μ , para os quais se verificam diferentes velocidades de convergência do algoritmo. Veja-se a Figura 4.20:

⁶ Essa actuação deverá estar dentro da gama dinâmica de resposta do igualizador – os $\pm 12\text{dB}$. De novo, não se procede à actualização de G_{est} através do cálculo de uma média aritmética simples, pelo problema ilustrado na Figura 4.9.

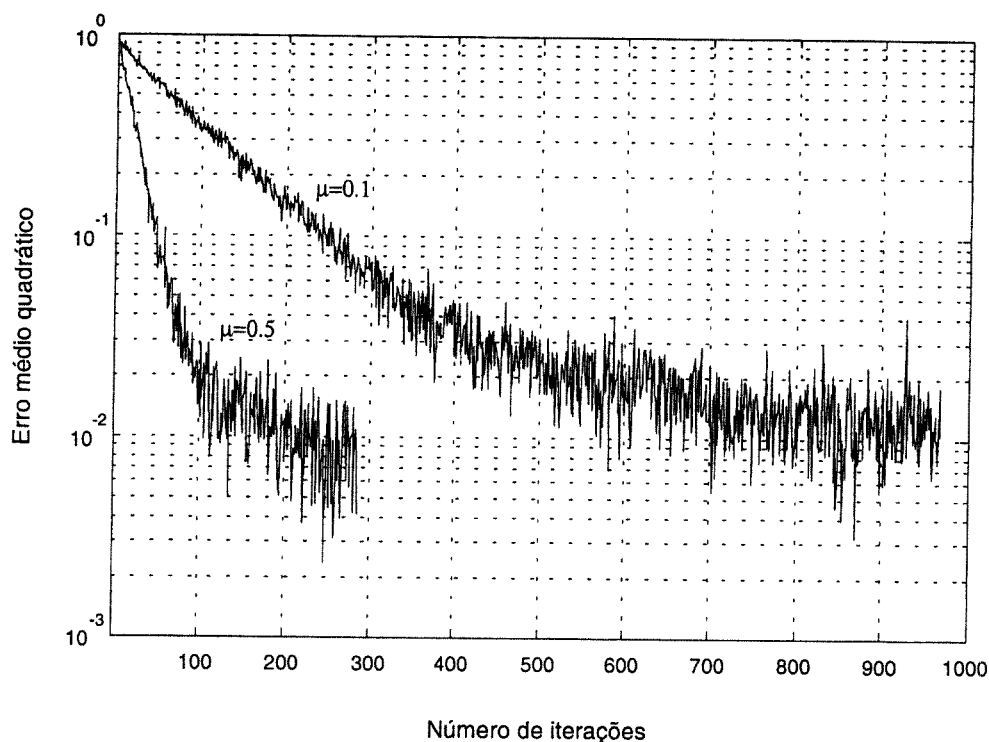


Figura 4.20 - Curva de aprendizagem para o método 4.2.4.3

Os valores de μ usados são 0.5 e 0.1 e na figura é representada a evolução do erro em função do número de iterações. Novamente, verifica-se uma maior rapidez de convergência do algoritmo para o passo de adaptação mais elevado. Comparando este erro com o obtido pelo método 4.3.2 é possível concluir que o método descrito neste capítulo apresenta um erro final inferior, o que pode estar associado à vantagem em utilizar sequências MLS como fonte de excitação, possuidoras de elevada SNR, uma vez que por possuírem um factor de *crest* mínimo, excitam o sistema com níveis de potência superiores. Adicionalmente, neste método obtém-se a resposta em frequência da sala através do cálculo de uma correlação cruzada de comprimento elevado, o que permite assumir que as últimas amostras dos sinais usadas no seu cálculo possuem já valores de energia estáveis. Por outras palavras, obtendo a resposta em frequência da sala através do cálculo da transformada de Fourier de um segmento considerável da parte mais significativa da sua resposta impulsional, obtêm-se respostas com pouca variância, o que para o caso de $\mu=0.5$ permite atingir um valor de erro estável (abaixo do qual este não descerá) ao fim de 250 iterações. Já para o passo de adaptação $\mu=0.1$, tal convergência surge apenas ao fim de cerca de 1000 iterações. Como os erros se aproximam, obtém-se mais rapidamente uma convergência aceitável usando um passo de adaptação de 0.5.

Veja-se a resposta após excitação inicial da sala (numa altura em que não houve ainda qualquer correcção) na Figura 4.21:

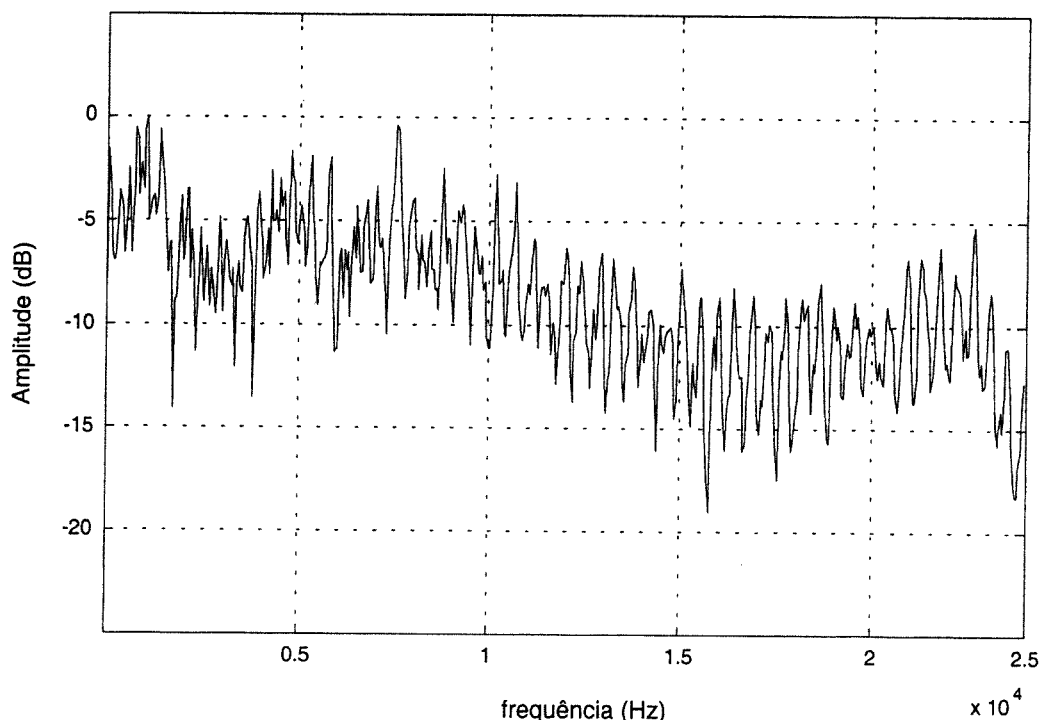


Figura 4.21 - Resposta inicial da sala para o método 4.2.4.3

É interessante comparar a figura de cima com a Figura 4.10. Por inspeção visual podem ser identificados os vales e picos mais relevantes, e a semelhança das suas localizações em ambos os espectros, principalmente para as baixas e médias frequências. Continuam presentes os vales a 2.5kHz e a 6kHz, enquanto que o centro do vale principal nas altas frequências parece estar agora ligeiramente abaixo, situado nos 15kHz. Pretende-se que estes vales e picos sejam anulados, para que o espectro seja corrigido para a desejada forma '*flat*'. As reflexões das ondas sonoras nas superfícies da sala provocam em determinadas frequências o efeito de reforço do sinal, enquanto que para outras o atenuam. Este efeito deve-se aos modos de ressonância da sala. Essas diferenças de amplitude entre frequências chegam a ser de quase 19dB, em alguns casos.

Através da excitação da sala por meio de sequências MLS e realizando o cálculo adaptativo dos novos coeficientes de igualização em cada iteração com base em (4.72), a partir do momento em que o erro (Figura 4.20) dá indicação de se ter atingido o funcionamento em regime estacionário, é possível visualizar os níveis de igualização (parâmetros do equalizador) automaticamente encontrados para realizar a correção do espectro da sala, para o ponto onde se encontra o ouvinte:

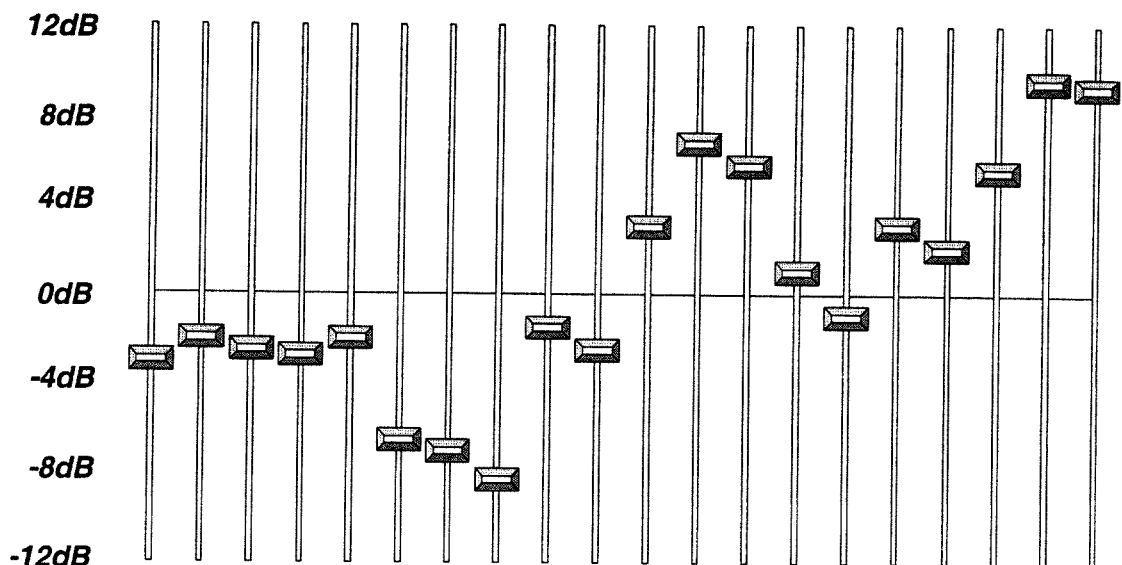


Figura 4.22 - Níveis de igualização da sala obtidos adaptativamente para o método 4.2.4.3

Compare-se a Figura 4.22 com a Figura 4.12 relativa ao método descrito na secção anterior 4.3.2.

BANDA	CENTRO (Hz)	NÍVEIS DE IGUALIZAÇÃO
1	47	-2.9
2	140	-1.8
3	234	-2.3
4	352	-2.6
5	492	-2.0
6	633	-6.3
7	820	-6.8
8	1008	-8.1
9	1242	-1.4
10	1523	-2.3
11	1875	3.1
12	2320	6.6
13	2930	5.6
14	3680	0.9
15	4805	-0.8
16	6352	2.8
17	8507	1.9
18	11414	5.4
19	15210	9.6
20	20883	9.2

Tabela 2 - Níveis de igualização para método MLS

Há uma semelhança evidente entre os perfis dos conjuntos de parâmetros de equalização encontrados pelos dois métodos. Com efeito, ambas as curvas de equalização possuem máximos e mínimos nas mesmas frequências, com valores bastante próximos. Vejam-se esses valores com mais detalhe na Tabela 2.

Desta forma, consegue-se a correcção do espectro em torno da média desejada. Comparando a Figura 4.23 com a Figura 4.21, é possível verificar esse ‘alisamento’ da resposta em frequência da sala. Compensam-se deste modo as colorações espectrais mais evidentes introduzidas pela amplificação, reprodução, propagação e captação do som.

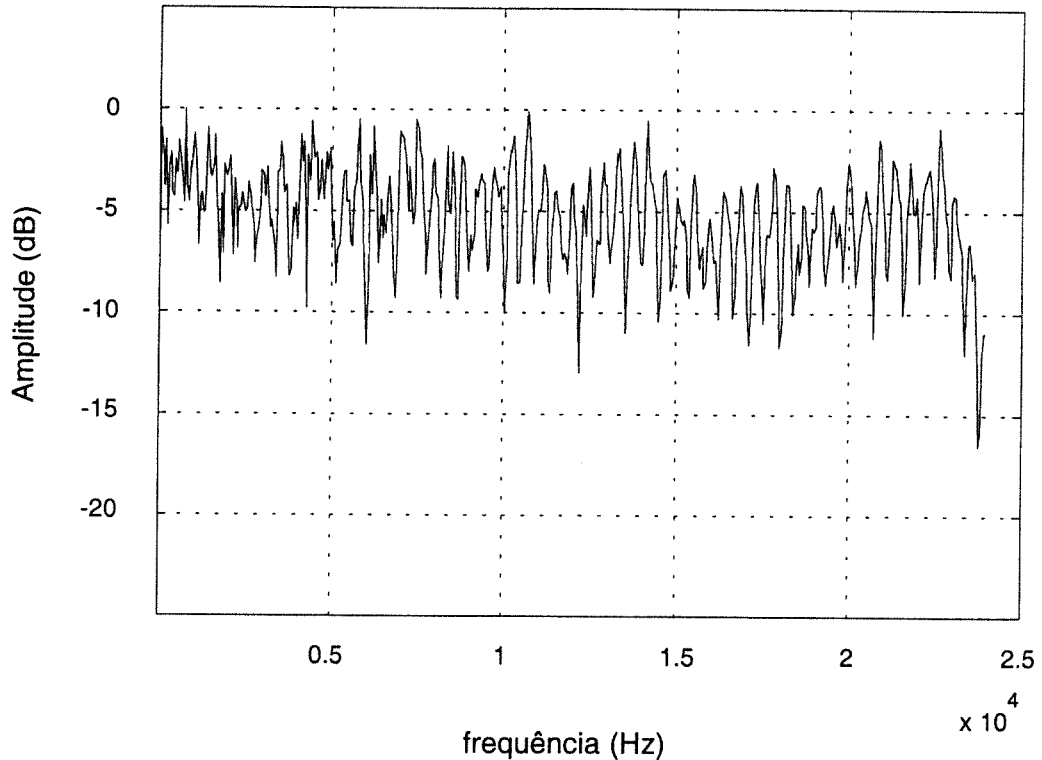


Figura 4.23 - resposta final da sala para o método 4.2.4.3

Como é visível, os vales foram suprimidos e os picos atenuados. O espectro fortemente recortado da Figura 4.23 deve-se à variância do sinal excitador e às inevitáveis reflexões naturais da sala. Observe-se a periodicidade do sinal que parece modular o espectro. Por inspecção visual, constata-se que surgem cerca de 11 máximos (períodos) desse sinal em cada banda de 5kHz. Observe-se agora na Figura 2.2 b), o atraso da primeira reflexão do sinal relativamente ao som directo. Analisando a transformada de Fourier de um sinal $x(n)$ e considerando uma única réplica desse mesmo sinal $\alpha x(n - n_0)$, atrasada de n_0 amostras e amortecida pela constante de absorção α :

$$\begin{aligned} x(n) &\longleftrightarrow X(e^{j\omega}) \\ \alpha x(n - n_0) &\longleftrightarrow \alpha e^{-j\omega n_0} X(e^{j\omega}) \end{aligned} \quad (4.73)$$

é possível demonstrar que a resposta em frequência da Figura 4.23, pode ser aproximada pela transformada de Fourier da sobreposição do sinal directo com a sua primeira reflexão. Com efeito, esta pode ser calculada por:

$$x(n) + \alpha \cdot x(n - n_0) \longleftrightarrow X(e^{j\omega}) \cdot (1 + \alpha \cdot e^{-j\omega n_0}), \quad (4.74)$$

cujas resposta em amplitude (módulo) pode ser expressa por:

$$\left| X(e^{j\omega}) \right| \sqrt{1 + \alpha^2 + 2\alpha \cos(\omega n_0)}. \quad (4.75)$$

Assim, para o atraso de cerca de 2.1ms da primeira reflexão relativamente ao som directo (Figura 2.2 b)), que à frequência de amostragem de 48kHz equivale a $n_0=101$ amostras de atraso, é possível constatar que a modulação provocada no espectro pelo cosseno da equação (4.75) é máxima, sempre que $\cos(\omega n_0)=1$. Considerando que a frequência angular normalizada ω varia entre 0 e π (i.e. varia entre 0 e a frequência de *Nyquist*), constata-se que neste intervalo é possível contabilizar 51 máximos, tal como se infere pela Figura 4.23, onde se pode contar em cada banda de 5kHz cerca de 11 máximos do cosseno modulador. Tais factos, explicam o espectro fortemente recortado da Figura 4.23.

De novo, mapeando este gráfico para as 20 bandas do sistema de equalização, obtém-se:

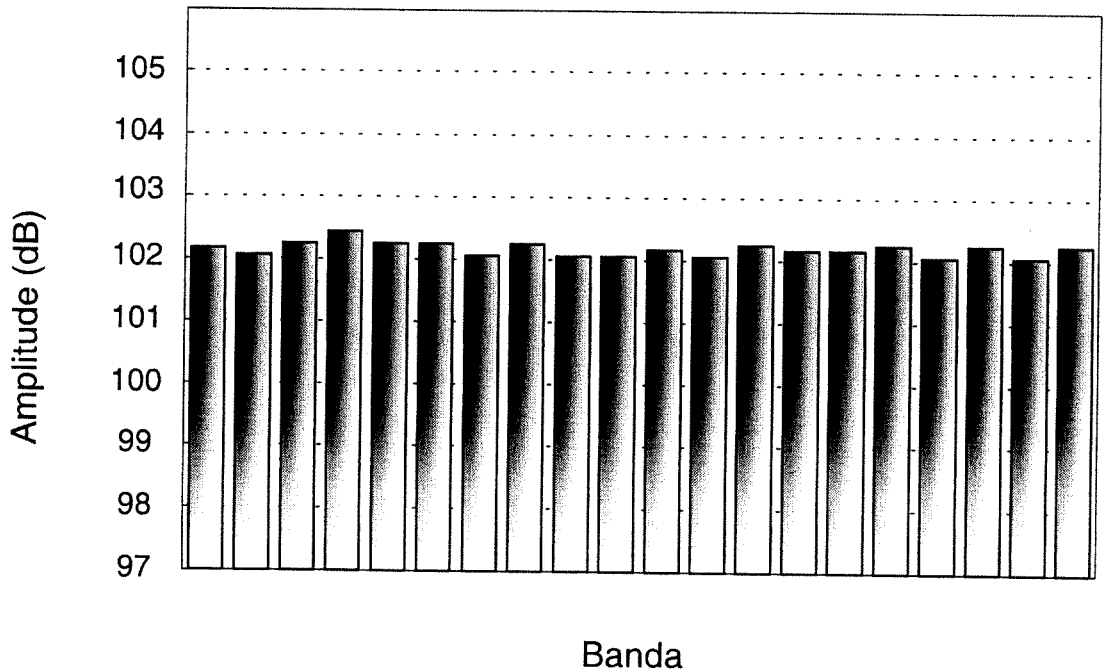


Figura 4.24 - Resposta em frequência nas 20 bandas para o método 4.2.4.3

Esta resposta em frequência (nas 20 bandas) mostra que o 'alisamento' espectral obtido não é perfeito, devido à presença de fenómenos acústicos na sala e à variância do sinal excitador. Adicionalmente, é sabido que sequências de ruído periódicas são extremamente vulneráveis a

efeitos de variação temporal [Müller&Massarini2001], muitas das vezes de natureza errática e carácter imprevisível.

O gráfico parece no entanto sugerir que com este método se obtêm resultados ligeiramente melhores (quando comparados com a Figura 4.14), como parece indicar a evolução do erro na Figura 4.20. Adicionalmente, comparando esta última (Figura 4.20) com a figura da evolução do erro equivalente, no método anterior – a Figura 4.11, verifica-se que o erro para este último método (MLS) em regime estacionário, possui menor variância.

Dois tipos de erro dominam a resposta impulsional: distorção e ruído. Dado um período de excitação suficientemente longo, a cauda da resposta impulsional (IR) contém apenas componentes de erro devidas aqueles dois factores. A amplitude de excitação com uma sequência MLS diz-se óptima, quando a energia da cauda é mínima, comparada com a energia da parte inicial da IR. Usando energia de excitação demasiado elevada, no erro da cauda domina a componente de distorção. Se por outro lado a energia for reduzida, domina a componente de ruído. Se o período da sequência MLS for muito mais longo do que o comprimento da IR do sistema, a imunidade ao ruído e distorção pode ser consideravelmente aumentada, eliminando a cauda da resposta impulsional, uma vez que a resposta impulsional linear tipicamente se encontra nas primeiras amostras da medida, enquanto que contribuições devidas ao erro se encontram uniformemente distribuídas pela IR [Rife&Vanderkooy89] e [Dunn&Hawksford93]. Assim, sendo a componente de distorção uniformemente espaçada na IR, a sua contribuição pode em grande parte ser removida de qualquer análise que envolva cálculos de energia. Distorções de pequena ordem não são, no entanto, distribuídas de forma uniforme, e resultam em imagens esporádicas da parte linear (inicial) da resposta impulsional. Podem no entanto ser identificadas, repetindo a medida usando uma sequência MLS diferente, mas de ordem idêntica (a posição dos artefactos depende da sequência utilizada [Vanderkooy94]). Após a realização desta operação, é possível identificar e remover esse tipo de distorções, fazendo com que a resposta impulsional seja uma representação mais fidedigna da componente linear da resposta impulsional do sistema [Vanderkooy94].

5. SISTEMA IGUALIZADOR

O bloco designado por '*Cálculo de novos parâmetros de igualização*' da Figura 4.6, recalcula em cada iteração novos níveis de igualização e actua em forma de realimentação na entrada de um outro bloco designado por '*Igualização adaptativa*', responsável pela igualização do sinal. Neste capítulo, pretende-se descrever o algoritmo de filtragem rápida por blocos implementado, que explora a vantagem oferecida pelo método de convolução rápida no domínio das frequências, designado por *Overlap-and-Add (OA)*, usando algoritmos de FFT (*Fast Fourier Transform*).

5.1. DIAGRAMA DE BLOCOS DO SISTEMA

Foi concebido um sistema de buffers para processar e armazenar as amostras áudio vindas da entrada, bem como das amostras a enviar para a saída. Foram também utilizadas várias tabelas, nomeadamente '*twiddle tables*', que aceleram o processamento das FFTs, já que dispensam a realização de alguns cálculos matemáticos. De facto, sabendo a dimensão da transformada, foi-nos possível calcular com antecedência os valores dessas tabelas, na figura representadas pelas exponenciais $e^{+j.k\frac{\pi}{N}}$ e $e^{-j.k\frac{\pi}{N}}$, e pelas janelas de análise e síntese. O diagrama de blocos do sistema para o canal esquerdo é apresentado na Figura 5.1 (o processo é em tudo análogo para o canal direito), desde a entrada do áudio até à sua saída devidamente igualizada.

Os números dentro dos círculos negros indicam a sequência temporal das operações realizadas, cuja descrição se apresenta de seguida. O parâmetro N utilizado para definir o tamanho dos buffers refere-se ao comprimento da FFT e tem, no caso específico deste projecto, o valor 1024.

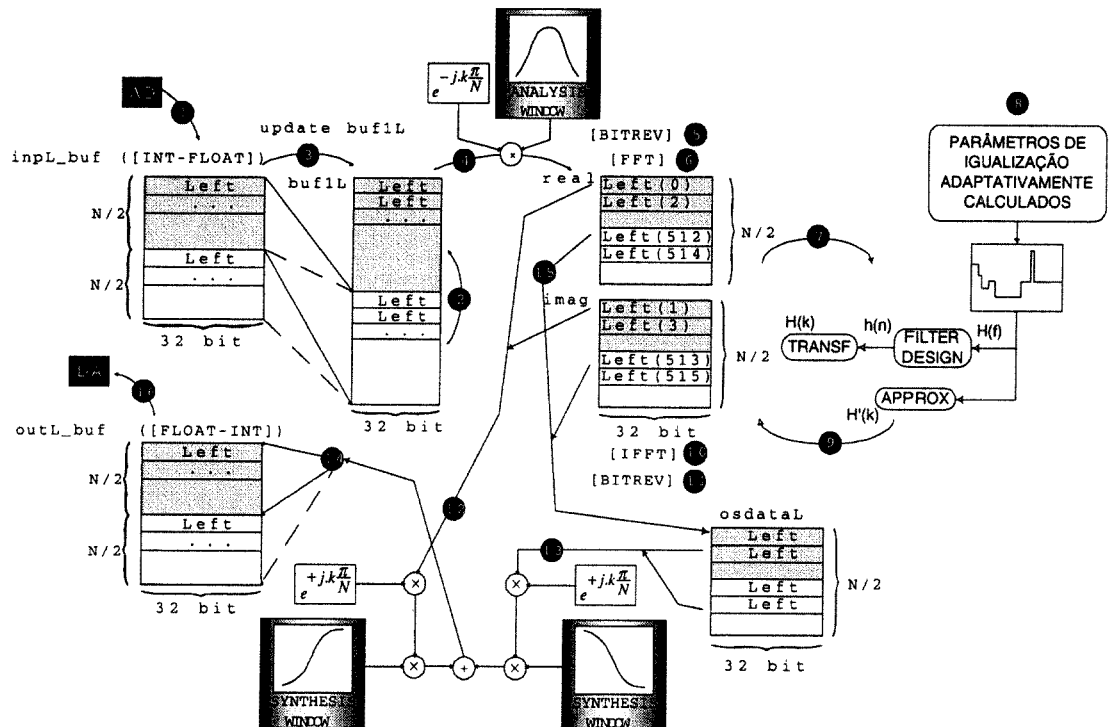


Figura 5.1 – Diagrama de blocos do sistema

Como referido acima, o processamento é análogo para o canal direito, motivo suficiente para que a explicação detalhada do algoritmo de equalização seja dada apenas para o canal esquerdo, descrito na figura. Seguindo desta forma a sequência temporal, temos que:

1. Dois *buffers* circulares, '*inpL_buf*' e '*outLbuf*', são usados para guardar, respectivamente, sinais áudio provenientes da entrada A/D do sistema e os correspondentes sinais processados a enviar para a saída D/A do mesmo. São blocos de memória endereçados de forma circular, de tamanho 1024, onde guardando ou enviando amostras a uma cadência de 48 KHz, se garante que o processamento de tramas áudio consecutivas é feito a uma velocidade superior à do preenchimento desses mesmos *buffers*, garantindo assim processamento em tempo-real. Inicialmente, os apontadores para estes dois *buffers* encontram-se na primeira posição de cada um deles. De cada vez que uma nova amostra chega, vinda do conversor A/D ou o D/A pede que lhe seja enviada outra amostra que converterá para a saída (passos 1 e 16), uma rotina de interrupção (ISR) é gerada. A Rotina de Atendimento a Interrupções (ISR) escreve uma nova amostra em '*inpL_buf*' e lê para a saída uma outra a partir de '*outL_buf*'. As amostras vindas do conversor possuem o formato 'inteiro' de 24 bits de comprimento e são convertidas para o formato 'float' de 32 bits, de forma a que se obtenha precisão nos cálculos realizados, aproveitando o facto de estarmos a trabalhar com um DSP de vírgula flutuante, isto é, que dispõe de elevada gama dinâmica.
2. A segunda metade do *buffer* buf1L (de tamanho N/2) é copiada para a primeira metade, de forma a que o algoritmo de *Overlap-and-Add* possa ser usado (a segunda metade de buf1L

actua como um *buffer* de dados de entrada previamente guardados pelo processamento anterior).

3. Cada um dos dois blocos '*inpL_buf*' e '*outLbuf*' de tamanho N acomoda dois blocos de tamanho $N/2$ de amostras áudio ordenadas temporalmente. Cada vez que o apontador alcança o meio ou o final de '*inpL_buf*', significa que há um novo bloco de tamanho $N/2$ pronto para ser processado. É nessa altura que se inicia o processamento das $N/2$ amostras mais recentes dessa metade do *buffer*, recém chegadas do A/D, através da sua cópia para a metade inferior de *buf1L*. Concatenando os dois referidos blocos em *buf1L*, forma-se uma trama de tamanho N pronta para ser processada. Enquanto esse processamento acontece, a rotina de interrupção ISR vai escrevendo/lendo novas amostras nos/dos referidos *buffers* '*inpL_buf*' e '*outLbuf*' (sem interromper o processador, já que essa transferência de amostras é feita por DMA), com uma periodicidade igual à frequência de amostragem (48 KHz), na outra metade (nas outras $N/2$ posições de memória) dos *buffers*. Desta forma garante-se que não há sobreposição de amostras.
4. A multiplicação das amostras por uma exponencial⁷, resultará no cálculo da ODFT. Adicionalmente, são também multiplicadas por uma janela de análise, processo necessário no método *Overlap-and-Add*.
5. Dá-se de seguida uma operação de *bit-reversing*, para ordenação dos dados no cálculo rápido da transformada. Esta operação é realizada '*in-place*'. Uma operação '*non in-place*' seria mais rápida, mas convém recordar que há limitações do ponto de vista de memória interna e os *buffers* de dados ocupam 1024 posições de memória, o que é um tamanho significativo, atendendo à memória disponível.
6. O cálculo da FFT DIT complexa radix-2, é igualmente realizado *in-place*, pelo motivo descrito anteriormente. Como é referido num capítulo posterior, esta rotina foi totalmente escrita na linguagem *Assembly* do *DSP SHARC 21065L*, daí advindo um ganho computacional de cerca de 5 vezes, relativamente à versão optimizada em linguagem C. A transformada de Fourier discreta de um sinal discreto real, tem parte real par e parte imaginária ímpar, ou seja, o cálculo de uma DFT para um sinal deste tipo apresenta redundância, já que conduz a um duplo cálculo dos mesmos valores. A FFT utilizada tira partido desta propriedade, uma vez que para um sinal discreto $x(n)$ real de comprimento N utiliza uma FFT de comprimento $N/2$. Para tal, parte $x(n)$ em dois sinais $g(n)$ e $h(n)$ de comprimento $N/2$ cada, da forma que se indica de seguida:

$$\begin{aligned} g(n) &= x(2n) \\ h(n) &= x(2n+1) \end{aligned} \quad (5.1)$$

para calcular simultaneamente as suas DFTs

⁷ Esta pré-multiplicação pode contudo ser evitada numa implementação optimizada da ODFT [Ferreira98].

$$\begin{aligned}
w(n) &= g(n) + jh(n) \\
G(k) &= \frac{W_R(k) + W_R((-k)) + j(W_I(k) - W_I((-k)))}{2} \\
H(k) &= \frac{W_I(k) + W_I((-k)) - j(W_R(k) - W_R((-k)))}{2}
\end{aligned} \quad (5.2)$$

de onde se pode recuperar $X(k)$ da seguinte forma

$$\begin{aligned}
X(k) &= G(k) + W_N^k H(k) \\
X\left(k + \frac{N}{2}\right) &= G(k) - W_N^k H(k)
\end{aligned} \quad (5.3)$$

Assim, e após a FFT, temos o espectro do sinal disponível em dois *buffers* – ‘*real*’ e ‘*imag*’.

7. A filtragem é realizada multiplicando os coeficientes do filtro que resultaram da curva de igualização definida pelo algoritmo adaptativo, pelo espectro destes dois sinais que, na realidade, representam o sinal original (comprimento N) dividido em dois (comprimento N/2).
8. Neste ponto do processamento, são fornecidos os novos parâmetros de igualização a serem usados no cálculo dos coeficientes do filtro. O algoritmo responsável pela actualização desses parâmetros é adaptativo e encontra-se descrito no capítulo 4.3.
9. Neste ponto, dá-se efectivamente a multiplicação dos coeficientes do filtro pelo espectro do sinal, após o que este se encontra igualizado pela nova curva.
10. O cálculo da IFFT DIF complexa radix-2, é realizado *in-place* à semelhança do que acontece com a FFT, pelo motivo descrito anteriormente. Também esta rotina foi totalmente escrita em *Assembly*, daí advindo um ganho computacional de cerca de 5 vezes, relativamente à versão optimizada em linguagem C.
11. Segue-se o re-ordenamento das amostras pela operação de *bit-reversing*.
12. Por estarmos a utilizar uma ODFT, em vez de uma DFT, impõe-se a multiplicação por um factor exponencial. De modo a respeitar o método de *Overlap-and-Add*, é também necessária uma multiplicação pela primeira metade da janela de síntese.
13. Da mesma forma, para respeitar o *Overlap-and-Add*, o bloco de dados igualizado no processamento anterior em *osdataL* deve ser multiplicado pelo mesmo factor exponencial e pela segunda metade da janela de síntese.
14. Os resultados dos pontos 12 e 13 somados, constituem a reconstrução perfeita do sinal igualizado e são guardados no *buffer* circular ‘*outLbuf*’, prontos a ser enviados para o conversor D/A. Da mesma forma que no passo 3, o apontador para ‘*outLbuf*’ dá-nos uma indicação de qual das metades do *buffer* está a ser utilizada pela saída, o que nos permite saber que devemos guardar o bloco de dados recém processado na outra metade.
15. Quanto à segunda metade dos dados calculados no processamento actual (pontos 5 a 11), deverá ser guardada em ‘*osdataL*’ de maneira a garantir o processamento seguinte, através do processo de *Overlap-and-Add*.

16. O sinal áudio igualizado é enviado para o conversor D/A a uma cadência de 48 kHz, por meio de uma rotina de interrupção (ISR). As amostras são enviadas para o conversor no formato 'inteiro' de 24 bits.

5.2. PARÂMETROS DE IGUALIZAÇÃO

Em igualização digital são normalmente utilizados Filtros FIR, por possuírem características de resposta em fase linear e portanto não introduzirem distorções provocadas por atraso de grupo. Tipicamente, equalizadores digitais usam os coeficientes da resposta impulsional dos filtros para realizar uma convolução linear com o sinal áudio no domínio dos tempos. O número de bandas deste tipo de equalizadores é normalmente reduzido (e.g. 10), uma vez que filtros com elevada selectividade espectral implicam a convolução do sinal com um número elevado de coeficientes, o que é sinónimo de uma carga computacional elevada, não realizável em certas aplicações.

Uma implementação mais eficiente da convolução linear pode ser alcançada através de técnicas no domínio das frequências, como os métodos de *Overlap-and-Add (OLA)* e *Overlap-and-Save (OLS)*. Contudo, devem ser tomadas diversas precauções para garantir que a convolução circular na verdade implementada por estes métodos corresponda, na prática, à convolução linear. Tal implica o desenho de uma nova aproximação ao filtro desejado, sempre que a resposta em frequência do equalizador for modificada, aproximação essa que deverá ter em consideração o comprimento máximo do filtro que não deverá ser excedido, de modo a evitar a introdução de componentes de 'aliasing' no sinal. Este passo de construção do filtro de igualização aproximado ao desejado, é bastante exigente do ponto de vista computacional.

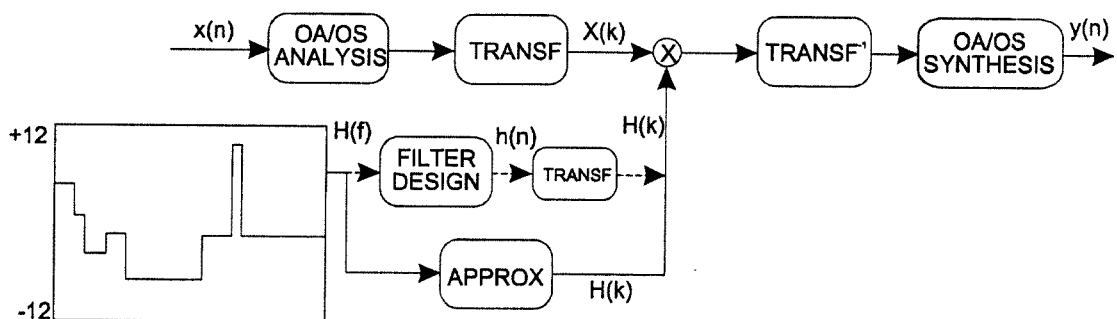


Figura 5.2 – Aproximação aos parâmetros de equalização desejados

Este é o aspecto mais crítico associado à filtragem eficiente no domínio das frequências. Devido às referidas limitações no seu desenho, o filtro de equalização final será sempre uma aproximação do desejado. Utilizando um algoritmo que calcule essa aproximação de forma rápida e eficiente, como é nosso objectivo (saltando portanto, por cima de um cálculo rigoroso e intensivo do filtro ideal), será de esperar o surgimento de componentes de 'aliasing' no sinal. Contudo, ao invés de procurar a eliminação matemática desse 'aliasing', basta-nos eliminar a sua perceptibilidade no áudio processado.

O diagrama de blocos da Figura 5.2 dá uma ideia da estruturação do algoritmo.

6. IMPLEMENTAÇÃO EM DSP EM TEMPO-REAL

6.1. HARDWARE

Porquê um Processador Digital de Sinal (DSP)?

O mundo, como o conhecemos, é basicamente analógico. E também o são os sinais que reflectem os processos que nele decorrem. Os computadores, por outro lado, são digitais. Dispositivos como conversores ADC proporcionam a conversão do sinal analógico para o formato digital enquanto que conversores DAC proporcionam a operação inversa. Uma vez neste formato, o processamento pode ser realizado por um DSP, capaz de cumprir essa tarefa de forma mais eficiente e a custo reduzido, quando comparado quer a uma solução baseada em tecnologia analógica, quer a um processador vulgar para PC. Os DSPs possuem características únicas que, ao contrário dos processadores para computadores pessoais, lhes permitem operar em tempo-real. Num PC comum, o utilizador habitualmente espera que o processamento seja concluído, enquanto que num DSP se obtém o resultado final sem se ter consciência do tempo gasto em processamento. Pode-se assim resumir as principais vantagens associadas ao uso de DSPs:

- Capaz de efectuar processamento complexo, impossível de realizar em circuitos analógicos;
- Algoritmos mais rápidos (microprocessadores e microcontroladores *correm* a uma frequência de relógio moderada e usam múltiplos ciclos de instruções para execução de um único passo. Um DSP possui instruções próprias para processamento de sinal que executam num único ciclo de relógio);
- Melhor qualidade do sinal e repetitividade de *performance* dos sistemas (uma vez que não dependem de componentes electrónicos, mas sim de um algoritmo que será sempre o mesmo);
- Mais flexíveis, uma vez que o processamento digital de sinal pode muitas das vezes ser facilmente modificado (e.g. actualização do algoritmo);
- Resulta normalmente numa aplicação de custo mais reduzido, para uma performance equivalente (redução no custo de material e no custo do trabalho). Menos componentes resultam em sistemas mais leves e pequenos.

Se a estas vantagens associarmos o facto de tipicamente possuírem uma série de periféricos ideais para a manipulação e controlo de dados (e.g. sinais áudio ou vídeo) como portas série, canais de DMA ou *timers*, compreende-se que são os processadores ideais para a implementação de algoritmos com aplicação na área de processamento de sinal em tempo-real.

6.1.1. ADSP SHARC 21065L ANALOG DEVICES (PROCESSAMENTO EM TEMPO-REAL)

O *ADSP 21065L SHARC (Super Harvard Architecture)* é um processador digital de sinal de arquitectura 32 bits e vírgula flutuante, com 544 Kbits de memória *on-chip*, concebido para uma vasta gama de aplicações, nomeadamente na área do áudio. O seu '*core*' de elevadas *performances* permite trabalhar a 180 MFLOPs (valor de pico), possui memória SRAM *on-chip* '*dual port*' (de acesso duplo), bem como periféricos I/O integrados suportados por um processador dedicado. Com a sua *cache* de instruções *on-chip*, é capaz de executar cada instrução num único ciclo de relógio. Quatro barramentos independentes de dados (de acesso duplo), de instruções e de I/O constituem a arquitectura do *ADSP-21065L SHARC*. Suporta as seguintes funcionalidades:

- Unidades de computação de vírgula flutuante de 32 bits (formato IEEE) – multiplicador, ALU e *Shifter* – que suportam 180MFLOPS;
- 16 registos de dados;
- Geradores de endereços de dados (DAG1 e DAG2);
- Sequenciador de programa com *cache* de instruções;
- 544 Kbits de memória SRAM '*dual port*' configurável;
- Controlador de DMA para suporte de 10 canais DMA;
- Duas portas série com dois registos receptores e transmissores cada, que suportam os modos TDM (*Time Domain Multiplexing*) e I²S;
- Dois *timers* programáveis e doze portas I/O programáveis e de uso genérico;
- 1 porta de acesso por JTAG.

O *21065L* possui dois barramentos PM (*Program Memory*) e DM (*Data Memory*), respectivamente de instruções e dados. O barramento PM pode aceder a instruções ou dados.

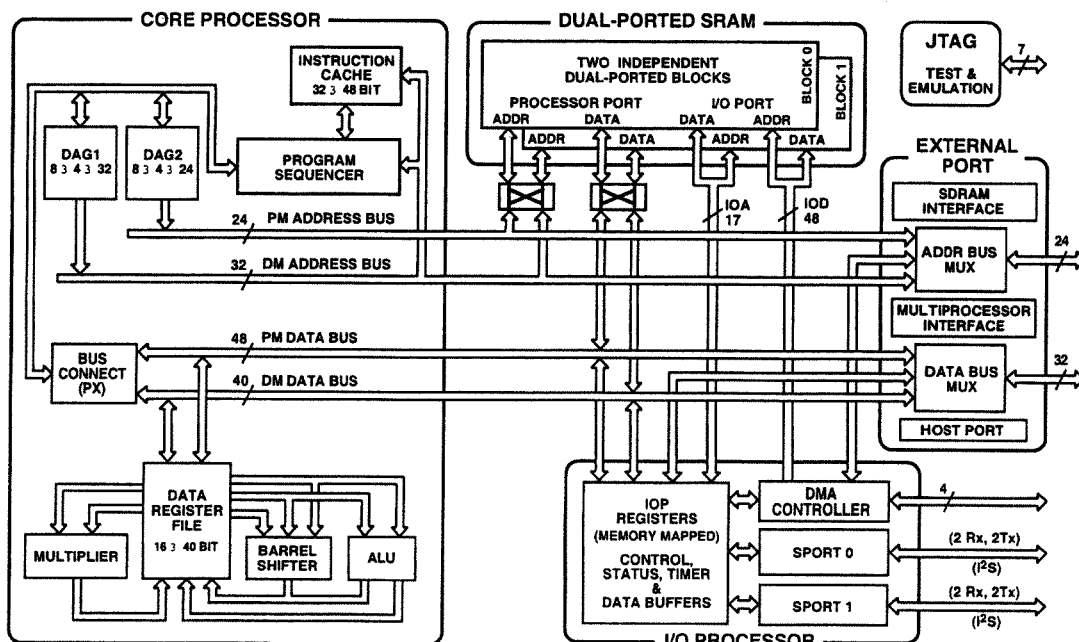


Figura 6.1 – Diagrama de blocos do ADSP-21065L Analog Devices

Durante um único ciclo de relógio, o processador pode aceder simultaneamente a dois operandos de dados – um através do barramento PM e o outro através do barramento DM, aceder a uma instrução da *cache* e realizar uma transferência por DMA.

Este SHARC possui os requisitos fundamentais para um DSP, de onde se destacam os principais:

Unidades de computação aritmética rápidas e flexíveis – O *ADSP-21065L* executa todas as suas instruções num único ciclo de relógio, proporciona ciclos rápidos e adicionalmente às tradicionais somas, adições multiplicações, e somas/multiplicações em paralelo, também proporciona um conjunto completo de outras operações aritméticas.

Circulação de dados de/para as unidades de computação sem estrangulamentos – Possui uma arquitectura eficiente, do tipo '*Super Harvard Architecture*'. Em cada ciclo, é capaz de:

- Ler ou escrever dois operandos de/para os registos;
- Fornecer dois operandos à ALU;
- Fornecer dois operandos à unidade de multiplicação;
- Receber dois resultados da ALU e da unidade de multiplicação.

A palavra de instrução de 48 bits suporta a transferência paralela de dados e operações aritméticas na mesma instrução.

Precisão adicional de 40 bits e elevada gama dinâmica – Suporta os formatos vírgula flutuante de 32-bit IEEE, 32-bit inteiro e fraccional (código complemento para dois e sem sinal), bem como o formato de precisão adicional de 40-bit vírgula flutuante IEEE. Este incremento de precisão, limita erros de truncatura nas suas unidades de computação. Os formatos de vírgula fixa possuem um acumulador de 80 bits que permitem o cálculo real de dados de 32-bit vírgula fixa.

Geradores de endereços duplos – Tem de igual modo dois geradores de endereços duplos (DAGs) que proporcionam endereçamento imediato ou indirecto com pré e pós-modificação. Suporta também operações de *bit-reversing* e endereçamento circular sem limitações no que diz respeito à zona de memória onde residem os buffers de dados.

Sequência de programação eficiente – Além de suportar o funcionamento em *loop* (ciclo) sem *overheads*, suporta também a entrada e saída desses mesmos *loops* num único ciclo. Os *loops* possuem até seis níveis de *nesting* (um *loop* diz-se *nested* quando existe dentro de outro *loop*), e podem ser interrompidos. Suporta de igual modo saltos com e sem atraso.

Ao nível do sistema, o *ADSP-21065L* inclui três grandes enriquecimentos:

Suporte de linguagens de alto nível – Várias características do ponto de vista da arquitectura que directamente permitem o suporte de compiladores de linguagens de alto nível e sistemas operativos:

- Registos de dados e endereços de uso genérico;
- Tipos de dados nativos de 32-bit;
- Grande espaço de memória endereçável;
- Endereçamento com pré e pós-modificação;
- Sem limitações no endereço de buffers de memória circulares;
- Programa, *loop* e *stack* de interrupção *on-chip*.

A sua arquitectura é concebida para suportar ANSI C *standard*.

JTAG para emulação – O *ADSP-21065L* suporta o *standard* IEEE P1149.1 *Join Test Action Group* (JTAG) para teste de sistemas. Este *standard* define um método de pesquisa série do *status* de cada componente do sistema (registos, memória, etc.). O emulador *ADSP-21065L EZ-ICE* usa de igual modo a porta série de emulação JTAG para aceder às funcionalidades de emulação *on-chip*.

Formatos IEEE – Suporta formatos de dados em vírgula flutuante do IEEE, o que significa que algoritmos desenvolvidos em processadores IEEE compatíveis, poderão ser transportáveis entre diferentes processadores, sem receio de instabilidade introduzida por arredondamentos, ou por manipulação de erros inconsistente.

E porquê usar um DSP de vírgula flutuante?

O formato de dados suportado por um processador de sinal determina a sua capacidade em lidar com sinais de variados tipos de precisão, gama dinâmica e SNR. São várias as vantagens de uma arquitectura vírgula flutuante:

Precisão – O número de bits de resolução dos conversores A/D tem vindo a aumentar, e a tendência é para que este aumento de precisão, bem como o da frequência de amostragem se mantenham.

Gama dinâmica – Um algoritmo trabalha tradicionalmente com sinais de precisão conhecida (e.g. 16 bit). A vantagem de um DSP com elevada gama dinâmica reside nos cálculos intermédios do processamento que, caso aquela não fosse elevada, poderiam excedê-la. A filtragem adaptativa é um tipo de aplicação que exige elevada gama dinâmica, o que vem reforçar a escolha do *ADSP-21065L* para a implementação desta aplicação.

Relação Sinal Ruído – Aplicações áudio, como é o caso da presente, exigem elevada gama dinâmica, de forma a permitir uma inquestionável separação de sinais que ocorrem em ambientes ruidosos.

Facilidade de utilização – De uma forma geral, DSPs de arquitectura vírgula flutuante 32-bit são mais fáceis de usar do que processadores de 16-bit vírgula fixa. A programação em linguagens de alto nível, uma vasta área endereçável e uma gama dinâmica elevada permitem que o programador se concentre no desenvolvimento do algoritmo em si, preocupando-se apenas com questões de algoritmia ao nível do processamento de sinal, e não com a dificuldade inerente à programação em *Assembly* e correcção de erros.

6.1.2. PLACA DE AQUISIÇÃO ÁUDIO AD1836

O AD1836 é um CODEC multi-canal da *Analog Devices*, capaz de trabalhar a uma frequência de amostragem de 96KHz e 24 bits de resolução. Possui dois conversores ADCs (entrada) e três DACs stereo (saída), de arquitectura *sigma-delta*. Possui também uma porta SPI incluída, permitindo a um micro-controlador ajustar e controlar diversos parâmetros (por exemplo, ajuste de volume ou monitorização de intensidade do sinal no ADC). Os conversores ADC apresentam 105 dB de SNR e gama dinâmica, enquanto que os DAC atingem os 108 dB para as

mesmas grandezas. Uma das grandes vantagens deste CODEC, reside no facto de possuir um modo de funcionamento com interface 'glueless' para uma porta série de DSP. Esse interface poderá funcionar no modo TDM (*Time Domain Multiplexing*), já que o dispositivo consegue empacotar até oito canais de entrada/saída numa trama série, fornecendo-os de seguida ao DSP, usando para o efeito uma única porta série do SHARC. A Figura 6.2 descreve o referido formato de empacotamento dos vários canais na forma e sequência com que chegam à porta série do DSP, que posteriormente se encarrega de os separar.

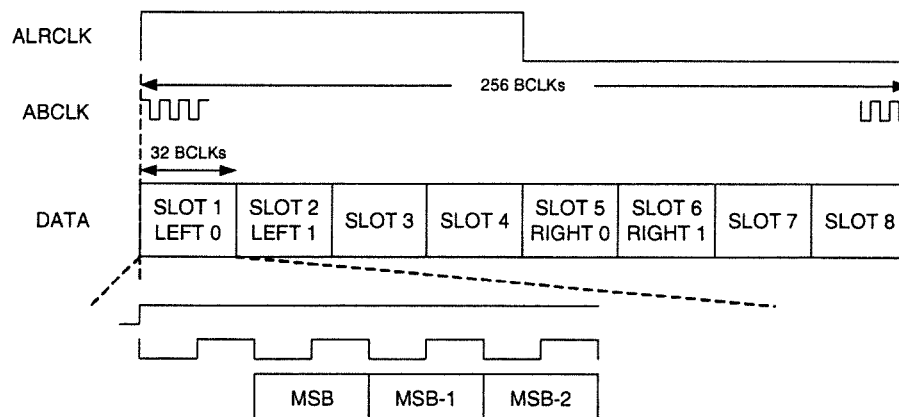


Figura 6.2 – Formato de trama do ADC

Todos os parâmetros que directamente influenciam o funcionamento do CODEC (resolução, frequência de amostragem, formato de empacotamento de dados, etc.) são programáveis por software, via porta SPI. O CODEC funciona como 'master', gerando portanto, todos os sinais de relógio que vão sincronizar o seu funcionamento com o DSP, ficando este último a funcionar em modo 'slave'. Neste modo de funcionamento, a 'board' do AD1836 fornece os sinais de trama (ALRCLK) e de relógio (ABCLK) ao processador, fazendo com que as respectivas conversões ADC e DAC fiquem temporalmente sincronizadas com as portas série do DSP, designadamente nas tarefas de recepção e transmissão de amostras.

6.1.3. SISTEMA DE AMPLIFICAÇÃO, MICROFONE E COLUNAS

Microfone TR40

O microfone seleccionado para captura da resposta sonora da sala consiste no TR40 da Audix, um microfone pré-polarizado capacitivo, especialmente concebido para testes, medidas e calibrações. Possui um diagrama polar omnidireccional e características de resposta em frequência extremamente linear, 'flat' e precisa, com variações de $\pm 1\text{dB}$ dos 20Hz aos 19KHz. É capaz de lidar com níveis de SPL extremamente elevados (superiores a 140 dB) sem apresentar distorção. Por possuir uma resposta em frequência 'flat', o TR40 pode ser utilizado na captura da resposta sonora da sala, fruto de uma determinada excitação.

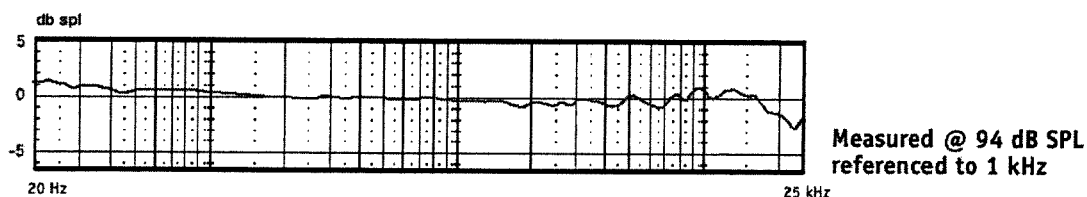


Figura 6.3 – Resposta em frequência do microfone TR40 Audix

Amplificador NAD C340

O sistema de amplificação usado baseia-se num amplificador *NAD* da série C 340. Apresenta uma resposta em frequência que varia entre $\pm 0.3\text{dB}$ na gama de 20Hz a 20KHz e SNR superior a 96dB.

Colunas B&W DM602 S2

Foram utilizadas colunas *B&W DM602 S2* de qualidade elevada, cuja resposta em frequência percorre a gama de 52Hz a 20KHz com uma variação na resposta em frequência de $\pm 3\text{dB}$.

Tendo em consideração o facto de as colunas introduzirem uma coloração (relativamente pouco significativa, atendendo à sua resposta em frequência) no espectro dos sinais excitadores da sala, será, à partida, de esperar que a resposta da sala captada pelo microfone venha influenciada por essa mesma coloração, o que de facto acontece. No entanto, não é necessário fazer uma análise detalhada do sistema para compreender que a igualização adaptativa feita à sala irá corrigir, não só as anomalias criadas por esta, relacionadas com reflexões do sinal, mas também a tal coloração (fixa) introduzida pelas colunas, uma vez que elas passam a fazer parte do canal de transmissão. Ora, como o objectivo deste trabalho é igualizar adaptativamente a sala, em busca de um espectro '*flat*' na gama analisada, fica claro que o sistema compensará as duas colorações introduzidas - a da sala (significativa) e a das colunas (pouco significativa) -, resultando daí a correcção desejada. Mais importante é, de facto, que o microfone possua uma resposta extremamente '*flat*', o que na realidade acontece, tendo em conta o cuidado tido na sua selecção, não descurando obviamente a escolha dos outros dois componentes, nomeadamente, amplificador e colunas.

6.2. SOFTWARE

O sistema implementado procurou encontrar uma forma de atenuar os efeitos provocados por reflexões de sinais acústicos e outras colorações espectrais induzidas pelos equipamentos amplificador e coluna numa sala. Com efeito, tais reflexões traduzem-se no reforço de determinadas frequências e atenuação de outras. O processo idealizado para efectuar essa compensação, consiste em igualizar a sala de forma adaptativa, de modo a minimizar o efeito das referidas reflexões. Para o efeito, é necessário utilizar um sinal conhecido possuidor de determinadas características, e com ele excitar a sala, após o que, obtida a resposta sonora a essa excitação, é possível retirar conclusões sobre o comportamento da sala. Descreve-se de

seguida, algumas das funções mais importantes utilizadas na implementação deste trabalho em plataforma DSP.

6.2.1. IMPLEMENTAÇÃO MIXED C/ASSEMBLY

O corpo principal deste programa (estrutura *main*) foi desenvolvido em C. Na realidade, para um observador exterior ao processo, todos os algoritmos parecem simplesmente implementados em linguagem C. As rotinas (computacionalmente mais intensivas) estão declaradas nesta linguagem mas, na verdade, definidas em *Assembly*, respeitando para o efeito a sintaxe das duas linguagens. Este método de implementação revelou-se prático, na medida em que numa primeira fase do trabalho se implementaram as rotinas intensivas do ponto de vista computacional em linguagem *Assembly*, e só depois se passou à fase de construção e teste dos algoritmos em si, tirando proveito dessas mesmas rotinas previamente implementadas e optimizadas, sem mais preocupações com questões da execução em tempo-real.

Várias das rotinas desenvolvidas na implementação deste trabalho têm utilização em mais do que um dos algoritmos testados. Com efeito, houve a preocupação de fazê-lo, na medida em que entre algoritmos diferentes varia por vezes o sinal excitador, ou o processo adaptativo em si, tal não impedindo que rotinas comuns sejam usadas. São disso exemplo, as rotinas de FFT, IFFT, ou as rotinas optimizadas de multiplicação de dados pelas '*twiddle tables*', etc.

As Transformadas Rápidas de Fourier directa e inversa, respectivamente **dirrealtransf()** e **invrealtransf()**, são rotinas comuns a todos os algoritmos e bastante optimizadas para o efeito, daí advindo grandes ganhos em tempo de execução, graças a uma cuidada implementação na linguagem *Assembly* do DSP. Outro exemplo é a rotina **spectrum()**, usada no cálculo do espectro do sinal, ou a rotina de desenho do filtro **calc_filter()** com base nos novos parâmetros de igualização.

Assim, passa-se a descrever as rotinas comuns aos vários algoritmos implementados. A maioria delas foi inicialmente implementada em linguagem C, pelo que após testes em *Assembly*, se tomou consciência dos enormes ganhos em termos computacionais daí advenientes. Com efeito, os ganhos de optimização em algumas foram superiores a cinco vezes.

buf1_update()

Esta rotina executa o passo 2 da Figura 5.1, na medida em que actualiza a primeira metade do buffer **buf1L** que vai conter o sinal áudio pronto para posteriormente ser multiplicado pela janela de síntese e convertido por meio de uma transformada rápida para o domínio das frequências.

copy_inpbuf_1st() e **copy_inpbuf_2nd()**

Ainda segundo a Figura 5.1, aqui é realizado o passo 3 através da chamada alternada a uma destas duas rotinas. As amostras mais actuais presentes na entrada do sistema, guardadas em **inpL_buf** são por intermédio destas rotinas copiadas para a segunda metade de **buf1L**.

Juntamente com a rotina anterior, este processo possibilita a utilização do método *Overlap-and-Save* com sobreposição de dados de 50%. De realçar que a simples passagem desta rotina de C para Assembly, acelerou o processamento por um factor de dez.

analysis_window()

Antes do cálculo da transformada rápida de Fourier, é necessária a multiplicação por uma janela de síntese (janela seno) a fim de evitar fugas espectrais da representação nas frequências. Equivale ao passo 4 da figura.

dirrealtransf()

A transformação do domínio tempo para o domínio frequência, dá-se no passo 6, não sem antes ser precedida pela operação de '*bit reversing*' (passo 5). Quer esta última, quer a transformada, são realizadas *in-place* por haver limitação na quantidade de memória interna disponível. Apesar desta contrariedade, a simples conversão do código em *Assembly* introduziu elevados ganhos em tempos de execução da rotina. Adicionalmente, foram exploradas as possibilidades de optimização de código inerentes às características do DSP, pelo que se obtiveram grandes ganhos adicionais de optimização dos tempos de execução. Esta era uma das primeiras metas a atingir na elaboração do trabalho, uma vez que no início do mesmo ficou claro que, sendo o tempo disponível para processamento de uma trama de 512 amostras em tempo-real de 10.67ms (frequência de amostragem 48kHz), as rotinas para cálculo das transformadas teriam forçosamente de ser implementadas no *Assembly* do DSP, por serem lentas quando programadas em C, o que comprometia o requisito de tempo-real. Com efeito, na sua versão C, a rotina executava em cerca de 3ms. Após passagem de código para a linguagem de baixo nível do processador, obtiveram-se tempos de execução da ordem dos 600µs, um ganho cinco vezes superior.

calc_filter()

Da teoria de processamento de sinal, é conhecida a propriedade que indica ser a convolução no domínio dos tempos equivalente à multiplicação no domínio das frequências. Como o processamento aqui efectuado se dá neste último caso, e como se controlam vinte bandas de igualização, impõe-se a adaptação dessas vinte bandas num vector descritivo do formato espectral que se pretende dar ao sinal áudio (igualização), de comprimento igual ao da transformada, de modo a que possa proceder à tal multiplicação nas frequências. Esta rotina faz o equivalente ao processo de selecção de novos coeficientes para filtragem (por convolução) no domínio dos tempos.

filter_data()

Fazendo uso dos coeficientes calculados na rotina anterior, aqui se procede à multiplicação dos novos coeficientes de igualização pelo espectro do sinal a igualizar. Esta rotina, assim como a descrita anteriormente, estão representadas na Figura 5.1 pelos passos 7, 8 e 9.

Invrealtransf()

Analogamente à transformada directa, a correspondente transformada inversa permite recuperar o sinal agora igualizado para o domínio dos tempos. Também aqui o processamento é *feito in-place*, toda a rotina se encontra implementada em *Assembly*, e foram igualmente exploradas as características do DSP que possibilitaram uma optimização muito significativa do tempo de execução. Uma vez mais, a transformada inversa executava em cerca de 3ms na sua versão em linguagem C. Após adaptação do código para o *Assembly* do processador, obteve-se um tempo de processamento cinco vezes mais rápido, também este da ordem dos 600µs. Posteriormente ao cálculo da transformada rápida, surge a reordenação temporal das amostras por '*bit reversing*' (passos 10 e 11).

copy_outbuf_1st() e copy_outbuf_2nd()

Estas duas funções são responsáveis pela reconstrução perfeita do sinal no domínio dos tempos, através do método *Overlap-and-Add*. Com a metade dos dados mais recentes e com a metade dos calculados no processamento da trama anterior, ambos multiplicados pela janela de síntese e adicionados, é pois possível recuperar o sinal com reconstrução perfeita⁸, de novo no domínio temporal. Posteriormente, é necessário decidir em qual das metades do buffer de saída outL_buf será armazenado o áudio acabado de processar. Por este motivo existem duas funções, sendo que cada uma delas é responsável pela cópia para um dos dois locais possíveis dentro do buffer. A Figura 5.1 representa estas operações pelos passos 12, 13 e 14.

osdata_update()

Na sequência do que foi dito anteriormente, o processamento de uma trama guarda uma parte do sinal para o processamento seguinte. Como descrito no passo 15 da figura, esta rotina faz essa operação de salvaguarda de dados, actualizando em cada novo processamento o buffer osdataL.

spectrum()

Outra ferramenta comum a todos os algoritmos implementados e de grande importância no processo de adaptatividade, é esta rotina, responsável pelo cálculo e representação da potência do sinal nas 20 bandas de controlo do algoritmo, por análise dos 512 bins em que se divide o seu espectro, após a transformada directa.

Todas as rotinas descritas neste capítulo foram optimizadas para a linguagem *Assembly* do *ADSP 21065L SHARC*, o que permitiu a obtenção de ganhos consideráveis em tempo de execução do algoritmo. Na verdade, sem estas optimizações, não teria sido possível implementar o trabalho nem realizar experiências neste DSP, o que obrigaria à utilização de um outro com velocidade de processamento superior, o que implicaria custos adicionais.

Uma vez que, como o próprio nome desta dissertação indica, o requisito tempo-real era fundamental para a realização deste trabalho, parece-nos que globalmente o empreendimento

⁸ Na ausência de modificação espectral.

foi bem sucedido, se considerarmos o grau de desafio que o objectivo implícito de optimização implicava à partida.

6.2.2. OPTIMIZAÇÃO DE CÓDIGO PARA OPERAÇÃO EM TEMPO-REAL

Desempenho do código

Foram usadas técnicas de filtragem rápida no domínio das frequências para implementar o igualizador que corrige a resposta acústica da sala. Esse tipo de técnicas baseia-se em transformadas rápidas de Fourier (FFT) que, mesmo optimizadas em linguagem C, executavam em quantidades de tempo consideráveis, logo eram bastante exigentes do ponto de vista computacional. De forma a conseguir processar tramas de 512 amostras $\times$ 2 (stereo) em menos de 10.7 ms (para conseguir operação em tempo-real, uma vez que essas amostras chegam a uma cadência de 48KHz), impôs-se a realização de um esforço adicional na optimização de todo o código. Foi necessário um conhecimento aprofundado das instruções e arquitectura do *DSP SHARC 21065L*, bem como das suas restrições e capacidade de paralelismo de instruções.

As rotinas de cálculo de FFT e IFFT (computacionalmente intensivas) são implementadas na memória interna do DSP para execução mais rápida.

Foi tido especial cuidado na manipulação da 'cache', pelo que também houve a preocupação de não produzir código ineficiente do ponto de vista de utilização da mesma, bem como de não gerar conflitos no 'pipeline'.

Outra vantagem explorada neste processador reside no facto de ele ser capaz de efectuar o acesso simultâneo a dois blocos de memória diferentes numa única instrução, usando os barramentos 'Program Memory' (PM) e 'Data Memory' (DM), operação conhecida por 'Dual Data Accesses'.

O 'core' deste DSP possui ainda um controlador de 'Direct Memory Access' (DMA) para independentemente aceder a dados no mesmo ciclo de instrução.

A placa de desenvolvimento possui 16M Words de memória SDRAM externa com zero 'wait-states' de tempo de acesso.

O controlador de DMA permite ao processador ou a um dispositivo externo especificarem operações de transferência de dados e regressar ao processamento normal enquanto o DMA se encarrega dessas mesmas transferências tipicamente demoradas, de forma independente e transparente.

Uma vez que os registos 'Data Address Generators' (DAG1 e DAG2) e 'Register File' (F0 – F15) existem em duplicado no 'core' do DSP, é-nos possível usar um modo de 'Context Switching' onde por um mecanismo de troca de apontadores, os registos são salvaguardados numa única instrução. Permite de um modo muito rápido fazer a salvaguarda de registos após o surgimento de uma 'Interrupt Service Routine' (ISR).

O uso de instruções de salto com atraso (*'Jump delayed instructions'*) é mais uma forma de aumentar a eficiência do código. De facto, o uso deste tipo de instruções permite preencher o *'pipeline'* com instruções a executar ainda antes do salto para uma zona diferente do código, ao invés de esvaziar o *'pipeline'*, aguardar pelo salto e executá-las apenas depois.

Ao longo do programa, realizaram-se várias manipulações de *'buffers'* de dados, de forma a minimizar as cópias dos mesmos. Com efeito, e uma vez que lidamos com *'buffers'* de tamanho 1024, essas cópias, que consomem bastantes recursos do DSP, foram inseridas noutras operações realizadas aos *'buffers'*, ou então implementadas por DMA, libertando desta forma o processador.

Foi dada particular atenção à execução paralela do maior número possível de instruções. Este tipo de arquitectura permite a execução simultânea de três tipos de operações:

- Dupla soma/subtracção;
- Paralelo de multiplicação/ALU;
- Paralelo de multiplicação e soma/subtracção.

Estas três combinações de operações podem ainda ser usadas em simultâneo com dois acessos a memória, um pelo barramento de *'Data Memory - DM bus'* e outro pelo de *'Program Memory - PM bus'*. Pode-se desta forma, paralelizar quatro operações num único ciclo de relógio, o que obriga o programador a ter um cuidado especial na disposição de dados e registos (o paralelismo de instruções impõe grandes restrições na utilização de registos), conseguindo-se grandes ganhos em termos de capacidade computacional. O ciclo de *'butterfly'* da FFT/IFFT é a parte mais intensiva do algoritmo, facto este que nos fez dedicar grande atenção à sua optimização. Com efeito, implementando esta parte do algoritmo com um grau máximo de paralelismo, os ganhos obtidos em termos computacionais foram consideráveis, traduzindo-se num enorme aumento do desempenho global do algoritmo.

```
LCNTR=m1,DO butterfly UNTIL LCE;
f14=f2*f1;
f9=f0*f5;
f13=f0*f1, f0=dm(i2,m6), f1=pm(i10,m14);
f6=f2*f5, f9=f9+f14, f2=dm(i5,m6), f5=pm(i11,m14);

f13=f13-f6, f6=dm(i4,m5);
f15=f6-f13, f7=dm(i3,m5);
f10=f7-f9, dm(-2,i2)=f15;
f15=f6+f13, dm(-2,i5)=f10;
f10=f7+f9, dm(i4,m6)=f15;
butterfly: dm(i3,m6)=f10;
```

Figura 6.4 – Butterfly da FFT optimizada para SHARC 21065L

A Figura 6.4 dá uma ideia do grau de paralelismo conseguido. Concretamente, na quinta linha de código, é visível a obtenção do grau de paralelismo máximo, já que se usam 4 instruções em

paralelo num único ciclo de relógio – um paralelo de multiplicação com soma, combinados com dois acessos em simultâneo à memória, um através do barramento '*Data Memory*' e outro por '*Program Memory*'.

6.2.3. ROTINA DE ATENDIMENTO A INTERRUPÇÕES

Nesta aplicação, o processador atende dois tipos de interrupções, ambas activadas pelas duas portas série existentes no SHARC. Uma das interrupções surge apenas no início do programa e diz respeito à programação da porta SPI do CODEC, que o configura no modo de funcionamento desejado. A porta série '0' (SPORT0) é programada com este fim. O CODEC possui quinze registos que são programados uma única vez através desta porta durante a inicialização do sistema, não voltando tal a acontecer.

Posteriormente, o processador configura a outra porta existente, a porta série '1' (SPORT1), para através de DMA, receber e enviar amostras do sinal áudio para os conversores A/D e D/A, respectivamente. Quer a porta série SPORT1 do DSP, quer o CODEC estão programados para funcionar em modo multicanal, o que significa que uma única porta série gere a entrada e saída de vários canais em simultâneo. Por outro lado, o CODEC possui dois conversores A/D *stereo* e três conversores D/A *stereo*, sendo capaz de os alinhar numa trama sequencial, de forma a serem recebidos/transmitidos correctamente e de uma só vez pela porta série do DSP. Todo este processo de transferência de amostras em ambos os sentidos é feito em sincronismo com a actualização dos apontadores dos anteriormente referidos *buffers* de entrada/saída '*inpL_buf*' e '*outLbuf*'.

O controlador de DMA encarrega-se da transferência efectiva das amostras áudio nos sentidos A/D → DSP e DSP → D/A. À cadência de 48 kHz, a ISR interrompe o processamento para se encarregar de actualizar os *buffers* de dados (de entrada e saída) com essas novas amostras, para que o DMA possa fazer as futuras transferências devidamente actualizadas e sem sobreposição de dados.

Diferentes rotinas de interrupção (ISRs) são activadas por sinais de interrupção diferentes. A associação do sinal com a respectiva rotina é feita na tabela de vectores de interrupção ('*interrupt vector table*') durante o processo de inicialização.

7. TESTES E RESULTADOS

7.1. REALIZAÇÃO DE TESTES EM SALA COM CONDIÇÕES ACÚSTICAS DESFAVORÁVEIS

Na realização desta dissertação foram apresentados vários métodos de excitação acústica de uma sala e de compensação das anomalias decorrentes das suas características. Desses métodos, dois foram seleccionados para implementação em DSP em tempo-real: um recorre a uma fonte externa de excitação por ruído branco; o outro usa o sistema DSP para produzir sequências MLS e explora as suas propriedades matemáticas.

No primeiro método consegue-se um compromisso aceitável entre velocidade de convergência do algoritmo e erro final. Com efeito, ao fim de 50 segundos é possível assumir que a sala encontrou um conjunto de níveis de igualização muito próximo do ideal, que lhe permite aplanar a resposta em frequência para uma determinada zona da sala, compensando desta forma anomalias criadas por reflexões do sinal nas suas superfícies e outros vários factores acumulados como sejam as características de amplificadores e altifalantes.

O segundo método é mais elegante do ponto de vista matemático. A sua implementação permitiu verificar a exequibilidade de inúmeros conceitos matemáticos tipicamente usados na manipulação de sinais, nomeadamente na área de processamento de sinal. Foi implementado um algoritmo gerador de sequências MLS, com as quais se excita a sala e por estudo da sua resposta acústica se compensa anomalias do som. Verificou-se que este método apresenta um erro de convergência final inferior ao obtido pelo primeiro método.

7.2. APRECIACÃO SUBJECTIVA

Do ponto de vista subjectivo, após convergência do algoritmo adaptativo, há a noção clara de que a igualização obtida pelos dois métodos descritos, produz uma mudança perceptível na coloração espectral do sinal musical que excita a sala. A coloração *preferida* pelo ouvido humano é frequentemente diferente da coloração plana, obtida com a realização deste trabalho. Em todo o caso, depois de efectuada a compensação, obtém-se um som agradável.

8. CONCLUSÕES

As técnicas utilizadas na realização deste trabalho permitiram implementar um modelo eficiente de filtragem adaptativa no domínio das frequências, onde é feita uma abordagem diferente dos métodos convencionais. Com efeito, usam-se técnicas de filtragem rápida a partir de algoritmos de FFT otimizados para a linguagem do DSP de vírgula flutuante utilizado. Igualmente, o erro que permite ao algoritmo controlar o processo de convergência adaptativa é, ao invés do que normalmente acontece, uma grandeza no domínio das frequências.

Os dois algoritmos implementados permitem efectuar a alteração da resposta em frequência de uma sala, compensando as anomalias provocadas por reflexões do sinal nas suas superfícies e outros factores acumulados. Mostram duas soluções alternativas possíveis para o mesmo problema. Num, é utilizado ruído branco gerado por uma fonte externa ao sistema com o qual se excita a sala, para posterior captura e correcção da sua resposta. No outro método excita-se a sala com sequências MLS geradas pelo próprio algoritmo, onde através do aproveitamento da suas propriedades matemáticas se consegue extrair a resposta daquela e efectuar compensações. São descritas conclusões relativas aos dois métodos e comparados os respectivos desempenhos.

A realização deste trabalho permitiu também explorar de forma exaustiva as potencialidades do DSP *SHARC 21065L*, nomeadamente ao nível de optimização da sua linguagem *Assembly*, o que possibilitou o desenvolvimento de algoritmos óptimos do ponto de vista de velocidade de processamento. Na opinião do autor, foi talvez esta a mais aliciante das tarefas, que a realização deste trabalho exigiu. Com efeito, conseguiram-se tempos de execução muito reduzidos, para rotinas que eram inicialmente muito intensivas do ponto de vista computacional. No caso da FFT, desenvolveu-se uma versão na linguagem *Assembly* do DSP mais rápida que a própria rotina optimizada e disponível nas bibliotecas do ambiente de desenvolvimento da *Analog Devices*. Tais ganhos, tornaram possível a realização deste trabalho nesta plataforma DSP e deixam em aberto a possibilidade de acrescentar código, tendo em vista um melhor funcionamento dos algoritmos ou a introdução de novas funcionalidades.

Trabalho futuro

Com a implementação descrita nesta dissertação, conseguiu-se criar uma plataforma robusta e estruturada, na qual será facilmente possível introduzir modificações capazes de melhorar o desempenho do sistema. Os algoritmos implementados possuem reserva de tempo de processamento, uma vez que usam rotinas que foram optimizadas para o efeito e que permitem acrescentar código ao já existente, mantendo a condição de operação em tempo-real.

Um dos objectivos futuros consiste em criar métodos de compensação da resposta acústica de uma sala durante uma actuação ou espectáculo. Para o efeito, pensa-se utilizar técnicas de mascaramento de sinusóides (tornando-as imperceptíveis ao ouvinte) em sinais musicais, excitando a sala com a mistura dos dois, e extraindo a resposta da sala com base naquelas.

Outro método idealizado consiste em utilizar técnicas de mascaramento das próprias sequências MLS no sinal musical e com base na resposta daquelas obter a resposta da sala, para a qual se deverão efectuar as compensações necessárias.

Adicionalmente, parece possível adaptar este sistema, para realizar medidas de coeficientes de absorção de determinados materiais a ondas sonoras, em câmaras anecóicas.

BIBLIOGRAFIA

- “ADSP-21065L SHARC User’s Manual”. Analog Devices, September 1998.
- “ADSP-21065L SHARC Technical reference”. Analog Devices, September 1998.
- [Abrantes98] Sílvia A. Abrantes. “Processamento Adaptativo de Sinais”. DEEC/FEUP, 1998.
- [Beranek90] Leo L. Beranek. “Acoustics”. Published for the Acoustical Society of America by the American Institute of Physics, 1990.
- [Beranek92] Leo L. Beranek. “Concert hall acoustics”. Journal Acoustical Soc. America 92(1), p.1-40, July 1992.
- [Bharitkar&Kyriakakis2001] Sunil Bharitkar and Chris Kyriakakis. “New Factors in Room Equalization Using a Fuzzy Logic Approach”. Presented at the 111th AES Convention, New York, December 2001.
- “C Compiler Guide & Reference for the ADSP-2106x Family DSPs”. Analog Devices, July 1998.
- [Carvalho&Morgado&Henrique97] António P. O. Carvalho, António E. J. Morgado and Luís Henrique. “Relationships between Subjective and Objective Acoustical Measures in Churches”. Journal of Building Acoustics, Vol.4, No.1, 1997.
- “C Compiler Guide & Reference for the ADSP-2106x Family DSPs”. Analog Devices, July 1998.
- [Cooley&Tukey65] J. W. Cooley and J. W. Tukey. “An algorithm for the machine calculation of complex Fourier series”. *Mathematics of Computation*, 19, 90, pp. 297-301, 1965.
- [Cremer75] L. Cremer. “The different distributions of the Audience”. Applied Acoustics, Vol.8, pp. 173-191, 1975.
- [Dunn&Hawksford93] Chris Dunn & Malcolm Omar Hawksford. “Distortion Immunity of MLS-Derived Impulse Response Measurements”. Journal Audio Eng. Soc., Vol.41, No.5, p. 314-335, May 1993.
- [Everest89] F. A. Everest. “The Master Handbook of Acoustics”. Tab Books, 1989.
- [Farina&Righini97] Angelo Farina and Fabrizio Righini. “Software Implementation of an MLS Analyzer, with Tools for Convolution, Auralization and Inverse Filtering.” Presented at the 103rd AES Convention, New York, 26-29 September 1997.
- [Fernandes99] Gabriel F. P. Fernandes, Luís G. P. M. Martins, Miguel F. M. Sousa, Filipe S. Pinto, Aníbal J. S. Ferreira. “Implementation of a new Method to Digital Audio Equalization”. Presented at the 106th AES Convention, Munich, May 1999.
- [Ferreira&Vieira95] Aníbal J. S. Ferreira and José M. N. Vieira. “An Efficient 20 Band Digital Audio Equalizer”. Presented at the 98th AES Convention, Paris, February 1995.

- [Ferreira98] Aníbal J. S. Ferreira. "Spectral Coding and Post-Processing of High Quality Audio". *PhD thesis*, Faculdade de Engenharia da Universidade do Porto – Portugal, 1998, <http://telecom.inescn.pt/doc/phd_en.html>.
- [Ferreira99] Aníbal J. S. Ferreira. "An Odd-DFT Based Approach to Time-Scale Expansion of Audio Signals". *IEEE Transactions on Speech and Audio Processing*, Vol.7, No.4, July 1999.
- [Ferreira2001a] Aníbal J. S. Ferreira. "Combined spectral envelope normalization and subtraction of sinusoidal components in the ODFT and MDCT frequency domains". Presented at the *2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 21-24 October 2001, New Paltz, New York.
- [Ferreira2001b] Aníbal J. S. Ferreira. "Accurate Estimation in the ODFT Domain of the Frequency, Phase and Magnitude of Stationary Sinusoids". Presented at the *2001 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, 21-24 October 2001, New Paltz, New York.
- [Haneda&Makino&Kaneda97] Yoichi Haneda, Shoji Makino and Yutaka Kaneda. "Multiple-Point Equalization of Room Transfer Functions by Using Common Acoustical Poles". *IEEE Transactions on Speech and Audio Processing*, Vol.5, No.4, July 1997.
- [Haykin96] Simon Haykin. "Adaptive Filter Theory". Prentice-Hall, 1996.
- [Hidaka&Beranek&Okano95] Takayuki Hidaka, Leo L. Beranek and Toshiyuki Okano. "Interaural cross-correlation, lateral fraction, and low – and high – frequency sound levels as measures of acoustical quality in concert halls". *Journal Acoustical Soc. America* 98(2), Pt.1, p.988-1007, August 1995.
- [Johansen&Rubak99] Lars G. Johansen and Per Rubak. "Target Functions and Processing Techniques in Digital Equalization Design". Presented at the 106th AES Convention, Munich, May 1999.
- [Lindsay73] R. B. Lindsay. "Acoustics: Historical and Philosophical Development". 1973.
- [McClellan&Schafer&Yoder98] James H. McClellan, Ronald W. Schafer and Mark A. Yoder. "DSP First: A Multimedia Approach". Prentice-Hall, 1998.
- [McGrath94] David S. McGrath. "A New Approach to Digital Audio Equalization". Presented at the 97th AES Convention, San Francisco, November 1994.
- [Mitra98] Sanjit K. Mitra. "Digital Signal Processing – A Computer-Based Approach". McGraw-Hill, 1998.
- [Müller&Massarini2001] Swen Muller and Paulo Massarini. "Transfer Function Measurement with Sweeps". *Journal Audio Eng. Soc.*, Vol.49, No.6, p. 443-471, June 2001.
- [Nielsen97] Johan L. Nielsen. "Improvement of Signal-to-Noise Ratio in Long-Term MLS Measurements with High-Level Nonstationary Disturbances". *Journal Audio Eng. Soc.*, Vol.45, No.12, p. 1063-1066, December 1997.
- [Oppenheim&Schafer99] Alan V. Oppenheim and Ronal W. Schafer. "Discrete-Time Signal Processing". Prentice-Hall, 1999.

- [Oppenheim&Schafer75] Alan V. Oppenheim and Ronald W. Schaffer (1975). "Digital Signal Processing". Englewood Cliffs, New Jersey: Prentice-Hall.
- [Peterson&Ziemer&Borth95] Roger L. Peterson, Rodger E. Ziemer and David E. Borth. "Introduction to Spread Spectrum Communications". Prentice-Hall, 1995.
- [Pierce81] Allan D. Pierce. "Acoustics: An Introduction to Its Physical Principles and Applications". Published by the Acoustical Society of America, 1981.
- [Restivo97] Francisco J. O. Restivo. "Processamento de Sinal". DEEC/FEUP, Setembro de 1997.
- [Rife&Vanderkooy89] Douglas D. Rife and John Vanderkooy. "Transfer-Function Measurement with Maximum-Length Sequences". Journal Audio Eng. Soc., Vol.37, No.6, p. 419-444, June 1989.
- [Rife92] Douglas D. Rife. "Modulation Transfer Function Measurement with Maximum-Length Sequences". Journal Audio Eng. Soc., Vol.40, No.10, p. 779-790, October 1992.
- [Sabine1900] W. C. Sabine. "Architectural Acoustics". Am. Arch. Building News, Vol.68, 1900.
- [Smith97] Steven W. Smith. "The Scientist and Engineer's Guide to Digital Signal Processing". California Technical Publishing, San Diego, California, 1997.
- [Svensson&Nielsen99] W. P. Svensson and J. L. Nielsen. "Errors in MLS Measurements Caused by Time Variance in Acoustic Systems". Journal Audio Eng. Soc., Vol.47, No.6, p. 907-927, November 2001.
- [Tronchin&Farina97] Lamberto Tronchin and Angelo Farina. "Acoustics of the Former Teatro "La Fenice" in Venice". Journal Audio Eng. Soc., Vol.45, No.12, p. 1051-1062, December 1997.
- [TUT2002] "Room Acoustics". Tampere University of Technology, Finland.
<<http://www.cs.tut.fi>>
- [Vanderkooy94] John Vanderkooy. "Aspects of MLS Measuring Systems". Journal Audio Eng. Soc., Vol.42, No.4, p. 219-231, April 1994.
- [Vieira96] José M. N. Vieira. "Digital Five Band Equaliser with Linear Phase". Presented at the 100th AES Convention, May, Copenhagen 1996.
- [Zölder97] Udo Zölder. "Digital Audio Signal Processing". John Wiley & Sons, 1997.

