



Universidade Federal do Paraná
Setor de Tecnologia
Departamento de Engenharia Elétrica
Disciplina: Processamento de Sinais Digitais

Rudolfo Augusto E. Rüncos

GRR20043644

Equalização Adaptativa para Ambientes

Curitiba
2007

O problema da distorção causada pelo ambiente e suas possíveis soluções

Situações como concertos musicais ao vivo, palestras em auditórios ou até a reprodução de música ou filmes em uma sala de estar apresentam um problema em comum: a distorção provocada pelo ambiente no som. Características do ambiente como sua geometria, volume e tipo de superfícies ocasionam fenômenos de reflexão, reverberação, atenuação de faixas de frequência entre outros que alteram as características originais do áudio.

Além dessa distorção ser diferente para cada ambiente, ela também varia ao longo do tempo para um mesmo ambiente. Mudanças na disposição de objetos dentro do ambiente, alteração do tipo de revestimento de paredes ou chão, o número de pessoas dentro do ambiente em questão ou até mudanças de temperatura são exemplos de alterações que podem influenciar na resposta em frequência do ambiente, ou seja, como o ambiente modifica o som original.

As técnicas de equalização adaptativa para ambientes existem para tentar resolver esse problema. Elas analisam a distorção causada pelo ambiente e geram filtros que compensam essas distorções tudo de forma automática. Existem técnicas que geram uma compensação inicial à resposta em frequência do ambiente e mantêm essa compensação inalterada durante toda a apresentação, e outras, mais recentes, que geram compensações que são atualizadas periodicamente durante uma apresentação.

No primeiro caso, a equalização adaptativa é conseguida pelo uso de sinais como o ruído branco ou sequências pseudo-aleatórias para excitar o ambiente. A análise da resposta do ambiente a esses sinais permite a geração do filtro que compense essas distorções. Esse método não é o mais adequado para equalizar um ambiente porque, conforme dito anteriormente, a resposta em frequência de um ambiente pode variar ao longo do tempo, e um filtro que compensa as distorções no início de um concerto de música, pode não mais servir para o final da apresentação.

O segundo caso já é bem mais adequado para a situação prática, pois se adapta em tempo-real às variações que o ambiente pode sofrer ao longo do tempo. Isso é interessante para manter a fidelidade do som ao longo de toda uma apresentação ou reprodução de áudio. Essa forma de equalização adaptativa será o assunto deste trabalho. Mais precisamente, este trabalho se baseia em dois trabalhos semelhantes, um realizado por Rieta e Sanchis [1], e outro por Ferreira e col. [2], para explicar como essa técnica de equalização adaptativa pode ser implementada.

Rieta e Sanchis [1] desenvolveram um método para realizar a equalização do ambiente utilizando o Matlab como principal ferramenta. Da forma como esse trabalho foi apresentado, ele não representa uma solução em tempo-real do problema. No entanto, seu método poderia ser implementado em um processador dedicado de DSP e funcionar em tempo real, eventualmente com algumas modificações. Portanto, esse trabalho foi escolhido como um exemplo de como poderia ser realizada a equalização adaptativa pelo uso transparente de tons de teste.

Já Ferreira e col. [2] implementaram sua solução num chip de DSP, mostrando como que esse método pode ser aplicado na prática em tempo-real. Seu trabalho se baseou num estudo anterior que mostrava a implementação de um sistema de equalização adaptativa baseado no uso de ruído branco.

É válido comentar que esse tipo de equalização funciona adequadamente em sistemas de áudio que apresentem uma baixa distorção. Na prática, no entanto, isto dificilmente será um fator impeditivo ao uso da técnica, pois as situações que utilizariam-se desse método muito frequentemente utilizam sistemas de áudio de alta fidelidade.

Equalização adaptativa de ambientes através do emprego de tons mascarados

Como a resposta em frequência de um ambiente varia ao longo do tempo, para que seja possível compensar corretamente a distorção sobre o áudio, é necessário conhecer essa resposta não apenas antes de uma reprodução de áudio, mas também durante essa reprodução. No entanto, na maioria das vezes não é desejável parar a reprodução para medir as características do ambiente. Por exemplo, durante um concerto musical não é possível parar a apresentação para aplicar um ruído branco ao ambiente e obter sua resposta em frequência. Nem é interessante se parar a reprodução de um filme para poder corrigir a distorção no áudio. É necessário que o sistema de equalização adaptativa possa realizar os ajustes necessários à compensação durante o acontecimento do evento.

Para tanto, uma solução possível a esse problema é inserir ao sinal original tons que não sejam audíveis, mas que possam ser recuperados através de um microfone de teste no ambiente onde o sinal é reproduzido. Comparando-se os tons inseridos inicialmente no sinal de áudio com aqueles capturados pelo microfone é possível estimar a resposta da sala e gerar um filtro que compense essa resposta, fazendo com que o som que seja ouvido seja o mais próximo possível do sinal original.

O efeito do mascaramento de frequências

Para entender como é possível inserir tons de forma transparente num sinal de áudio, é necessário entender uma característica do ouvido humano que é o mascaramento de frequências. O mascaramento de frequências ocorre quando um som em uma dada frequência impede que um segundo som em uma frequência próxima seja ouvido. Por exemplo, um sinal de áudio que num dado instante apresente um pico de potência numa frequência "x" específica e uma menor potência em uma frequência "y" próxima a essa é ouvido como contendo apenas a componente de frequência "x". Diz-se que o som da frequência "x" mascarou o som na frequência "y".

Esse fenômeno de mascaramento pode ser entendido mais precisamente observando-se o gráfico da **figura 1**. Esse gráfico mostra o nível de mascaramento em dB provocado por uma frequência mascaradora f_M . O nível de potência de f_M é dado pela altura do retângulo vazado. PMR, do inglês *Probe-to-Masker Ratio*, representa o limite do nível de potência para uma frequência f_S , distante df de f_M , abaixo do qual f_S será mascarada por f_M . Esse limite é medido em dB abaixo do nível de potência de f_M .

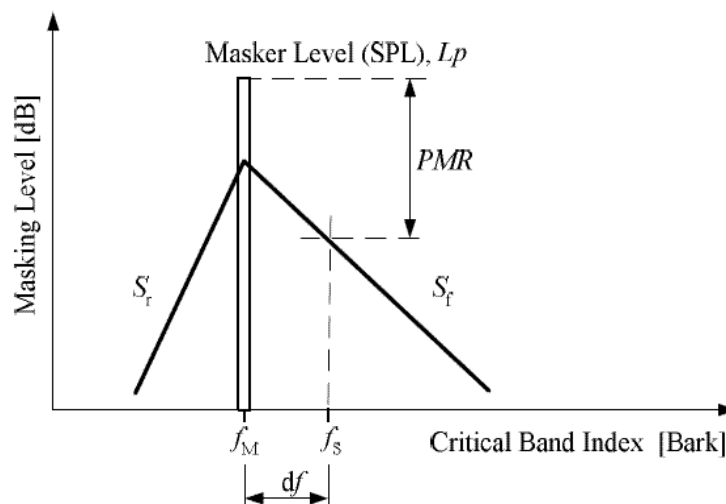


Figura 1: Curva do limite de mascaramento causada por um tom em f_M . Fonte: [2]

De acordo com o gráfico, para frequências abaixo de f_M o valor de PMR aumenta mais rapidamente do que para valores superiores a f_M . Ou seja, o efeito de mascaramento de frequências abrange uma maior banda de frequências superiores a f_M do que inferiores.

Análise do sinal original e inserção de tons de teste

Para que seja possível inserir tons no sinal de áudio de forma que estes sejam mascarados, é feita uma análise do sinal original. Essa análise é feita na distribuição espectral de potência do sinal original, de forma a reconhecer picos de potência (frequências mascaradoras) e faixas de frequência que estejam mascaradas por esses picos.

O sinal é analisado a curtos intervalos de cada vez. Rieta e Sanchis [1] realizaram a análise a cada 4096 amostras de seu sinal original amostrado a 44,1kHz (aprox. 0,093s), e Ferreira e col. [2] utilizavam 1024 amostras de seu sinal amostrado a 48kHz (aprox. 0,021s). Existem dois motivos para que a análise do sinal seja feita a curtos intervalos. O primeiro é diminuir o atraso causado ao sinal original. E o segundo é facilitar sua implementação, visto que há limitação da capacidade de memória e processamento para o número de amostras analisados por vez. No entanto, esse número deve ser suficiente para proporcionar uma boa resolução espectral do sinal de forma a permitir a análise e o correto posicionamento dos tons de teste.

Após identificadas as faixas de frequência que estejam mascaradas por picos de grande intensidade, o sistema deve decidir qual frequência dentro da faixa mascarada é a mais adequada para a inserção de um tom. Essa frequência deve ser escolhida de forma que haja a maior diferença de amplitude para frequências vizinhas, buscando assim minimizar a influência destas no tom de teste. A escolha da frequência também está limitada à resolução do espectro. A **figura 2** ilustra a análise das faixas e a escolha dos pontos de inserção de tons.

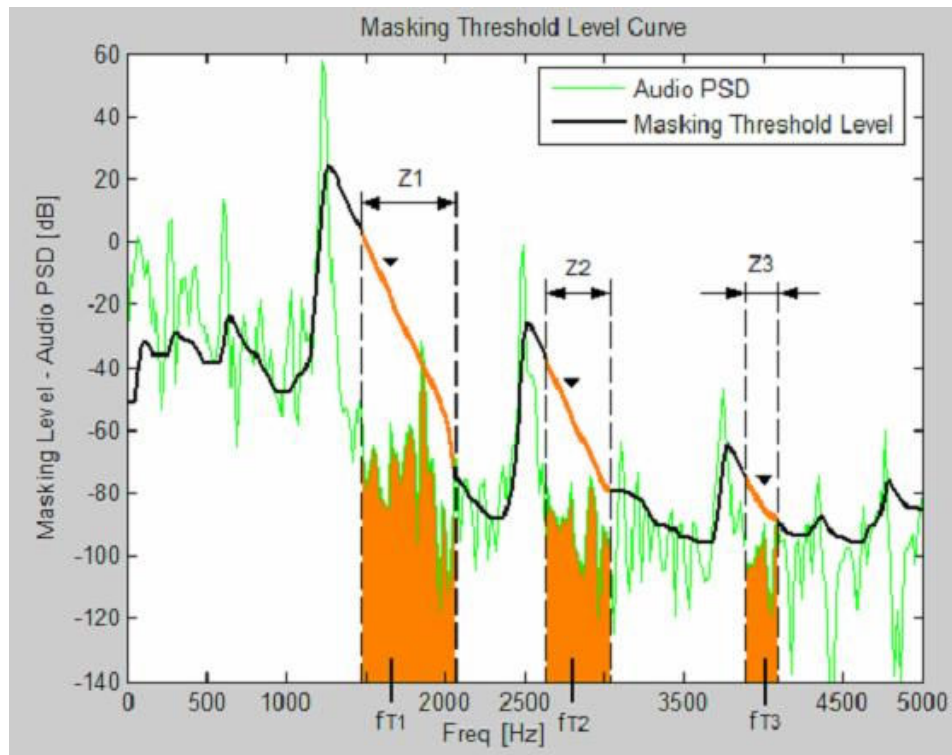


Figura 2: Distribuição espectral de potência para um dado intervalo do sinal original (em verde). A linha cheia em preto/laranja indica o limite para o efeito de mascaramento deste sinal. As regiões sombreadas em laranja mostram as faixas adequadas para a inserção dos tons de teste (Z1, Z2 e Z3). Os pontos marcados no eixo das frequências indicam os pontos escolhidos pelo sistema para a inserção dos tons de teste (f_{T1} , f_{T2} e f_{T3}).

Fonte: [2]

Com os pontos de inserção dos tons escolhidos, segue o processo de inserção. Antes de adicionar os tons, o sinal original passa por um filtro *notch* que rejeita com grande seletividade as frequências escolhidas para as inserções. Após feitos esses estreitos “buracos” no espectro exatamente sobre as frequências escolhidas, os tons de teste são somados ao sinal filtrado de forma a gerar o sinal de saída que irá para os auto-falantes. A **figura 3** mostra um exemplo desse processo de inserção do tom na frequência de 315Hz.

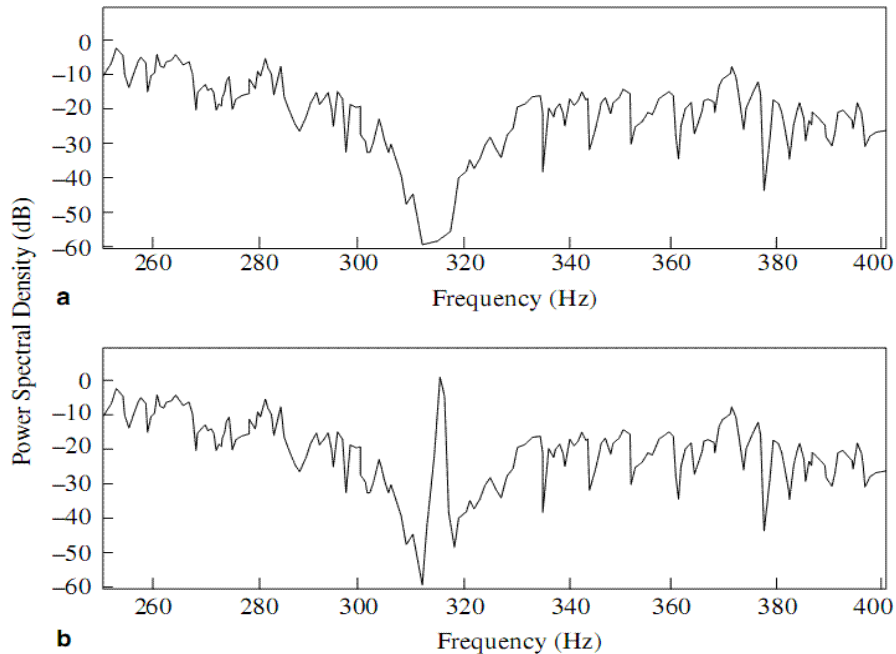


Figura 3: (a) Sinal resultante após a filtragem pelo filtro *notch*. (b) Sinal resultante após a inserção do tom.
Fonte: [1]

A intensidade e duração dos tons deve ser calculada ou escolhida de forma a mantê-los inaudíveis. É desejável que a intensidade seja a mais alta possível para garantir uma boa recepção do tom pelo microfone de teste. A duração é interessante que seja suficiente para garantir que todos os efeitos de reverberação e reflexão do ambiente atuem sobre o tom. Nesse sentido, ambos os autores [1 e 2] calculam a amplitude do tom de acordo com as características de potência do sinal original e dos limites de mascaramento para obterem um nível ótimo de amplitude. Já em relação à duração, Rieta e Sanchis [1] estipulam um tempo fixo, usualmente de 1s, e atualizam a amplitude do tom no início e no meio do tempo de duração do tom. No algoritmo implementado por Ferreira e col. [2], o tom dura o número de intervalos de análise consecutivos que permitirem sua inserção.

Estimação da resposta em frequência do ambiente e compensação no sinal

Os tons de teste inseridos no sinal são reproduzidos pelos auto-falantes e sofrem todos os efeitos do ambiente. O microfone de teste captura o sinal de áudio do ambiente, que contém os tons de teste cujas amplitudes foram alteradas pela resposta em frequência desse ambiente. O sistema de equalização mantém a informação de quais tons foram adicionados no sinal e quando, de forma a poder fazer a comparação. Para realizar essa comparação devem ser levados em conta os tempos de propagação do som dos auto-falantes até o microfone e eventualmente o tempo para que a resposta do ambiente ao tom estabilize. Com isso é possível comparar a atenuação (ou amplificação) sofrida por cada tom individualmente.

Para deixar o sistema independente de elementos de amplificação (amplificadores de potência) um modo de realizar a comparação é analisando a relação entre a potência do tom de

teste e a potência total do sinal tanto na saída (sinal que vai para a cadeia de amplificação de som) quanto na entrada do microfone de teste e fazer sua relação. Essa comparação é melhor descrita pela **equação 1** a seguir:

$$Att_T = P_{T_o} - G_o - (P_{T_r} - G_r) \quad (\text{equação 1})$$

onde Att_T é a atenuação sofrida pelo tom de teste T, P_{T_o} é a potência do tom T que foi inserido no sinal, G_o é a potência total do sinal de saída do sistema (que vai para os auto-falantes), P_{T_r} é a potência do tom recebido pelo microfone de teste e G_r é a potência do sinal recebido. Os valores dos termos desta equação devem estar todos em dB. O valor de Att_T indica o efeito que o ambiente tem sobre aquela frequência (em dB). Do resultado obtido por essa comparação para todos os tons inseridos no sinal é possível determinar parâmetros para a resposta em frequência do ambiente. Portanto, é possível também determinar os parâmetros para o inverso dessa resposta em frequência, ou seja, os parâmetros do filtro compensador.

Cabe ressaltar que o filtro compensador, que é o filtro que realmente atuará no sinal de entrada para compensar as distorções do ambiente, trabalha com um número de sub-bandas de frequência que é, via de regra, menor que o número de pontos que o sistema dispõe para representar o espectro dos sinais. Por exemplo, O sistema de Ferreira e col. [2] utiliza um filtro compensador de 16 sub-bandas, mas o espectro de seus sinais apresenta uma resolução de 512 pontos. Portanto, pode ocorrer que mais de um tom esteja dentro de uma mesma sub-banda desse filtro. Neste caso, o parâmetro que apresentar mais de um tom inserido em sua sub-banda é determinado utilizando-se a média das atenuações dos tons contidos nessa sub-banda. Para se obter um valor mais estável da resposta do ambiente, é aplicada também uma média temporal dos parâmetros calculados para cada sub-banda.

O filtro compensador é então atualizado com os diferentes parâmetros calculados para cada sub-banda de forma a compensar a resposta em frequência do ambiente. No trabalho de Ferreira e col. [2], esses parâmetros são atualizados de uma a duas vezes a cada 10 segundos. As chances de que nesse intervalo de tempo ocorra pelo menos uma medida para cada sub-banda são grandes. O autor considera que a resposta em frequência do ambiente se mantém constante nesse intervalo de tempo.

Um esquemático simplificado de um equalizador adaptativo de ambientes

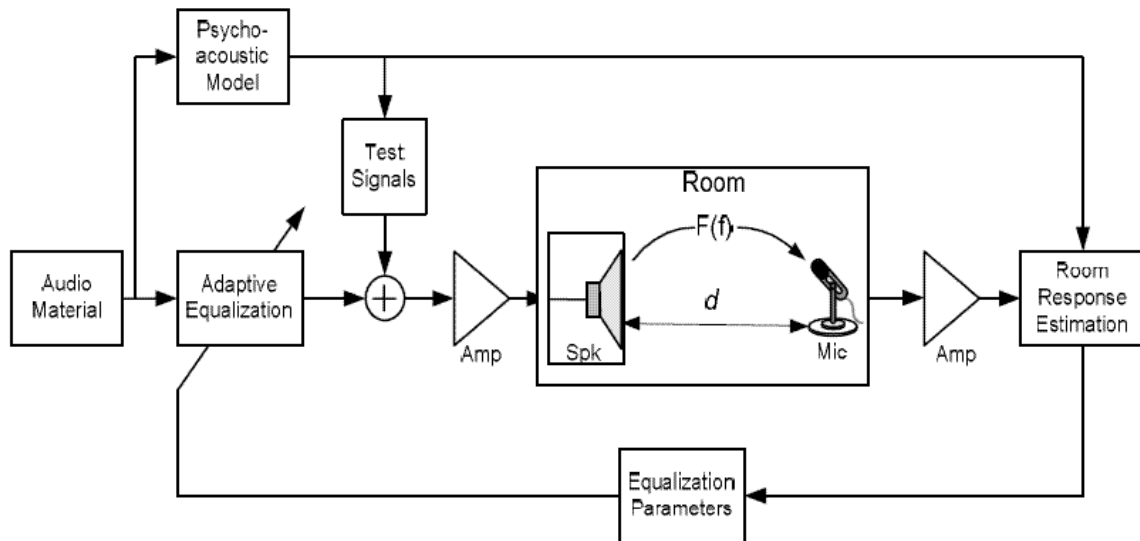


Figura 4: Diagrama em blocos simplificado de um equalizador adaptativo de ambientes. Fonte: [2]

A **figura 4** apresenta um diagrama em blocos simplificado de um equalizador adaptativo de ambientes. De acordo com esse diagrama, o sinal original (*Audio Material*) segue para o bloco que identifica as oportunidades para inserção de tons de teste (*Psychoacoustic Model*). Esse indica quais tons devem ser gerados (*Test Signals*) e inseridos no sinal que já passou pelo filtro compensador (*Adaptive Equalization*). O sinal segue para os amplificadores de potência que acionam os auto-falantes. O sinal, após sofrer as distorções do ambiente ($F(f)$), é capturado pelo microfone e devidamente amplificado. O resultado é utilizado para efetuar a comparação dos tons recebidos com os tons que foram inseridos segundo a saída do bloco que identificou as oportunidades. Dessa comparação (*Room Response Estimation*) sai a estimativa da resposta em frequência da sala, que é utilizada para gerar os parâmetros do filtro compensador (*Equalization Parameters*) que são utilizados para filtrar o sinal original de forma a compensar as distorções do ambiente.

Nesse diagrama não consta a filtragem que o sinal original sofre pelos filtros *notch* que criam um “buraco” para a inserção do tom. É interessante reparar que os tons são inseridos no sinal após ele já ter passado pelo filtro compensador. Isso porque deve atuar sobre os tons apenas a resposta em frequência do ambiente para que seja possível estimar esta.

Resultados e comparação com métodos de sinal interrompido

Rieta e Sanchis [1] compararam o resultado obtido pelo seu método com o resultado obtido a partir de um método de equalização de ambientes baseado na interrupção do sinal para aplicação de ruído rosa ao ambiente de forma a obter os parâmetros do filtro compensador. Essa comparação é mostrada na **figura 5**.

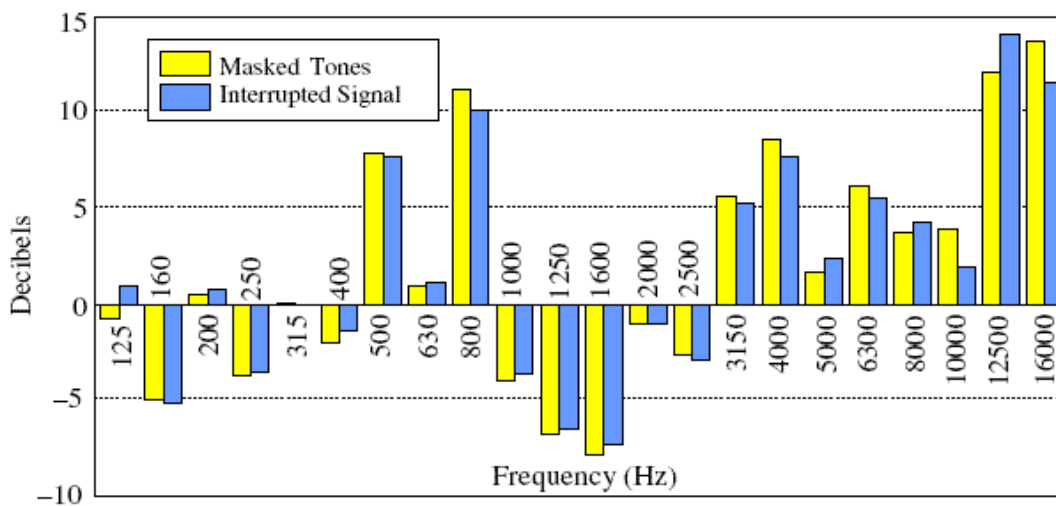


Figura 5: Comparação entre o método pela inserção de tons mascarados e o método de interrupção do sinal (ruído rosa) para a obtenção dos parâmetros do filtro compensador. Fonte: [1]

Nesse gráfico são mostradas os parâmetros obtidos para cada sub-banda (22 nesse caso) com centro nas frequências indicadas no eixo das frequências. Como os valores dos parâmetros são bem próximos, podemos considerar que a margem de erro dos dois métodos é similar.

Esse autor também fez um teste subjetivo com 25 pessoas sobre a diferença do som original para o som com os tons mascarados e a audibilidade destes. O resultado do teste é que não há diferença entre os dois sons pois os tons não são ouvidos pelas pessoas.

Ferreira e col. [2] também realizaram testes comparativos entre o seu sistema e outras técnicas mais tradicionais de obtenção da resposta do ambiente. Através dessa comparação foi concluído que os dois métodos obtêm resultados muito semelhantes, e que são portanto equivalentes em termos da capacidade de obter a resposta do ambiente. Ele [2] também afirma que os tons de teste inseridos pelo seu sistema são imperceptíveis mesmo para ouvintes “experts”. Ou seja, não há alteração do áudio que seja perceptível pelo ouvido humano.

Conclusões e perspectivas de desenvolvimento futuro

Os resultados deixam claro que esse método de equalização adaptativa de ambientes através da inserção de tons mascarados pode ser utilizado para compensar a resposta em frequência de um ambiente. Esse método é capaz de atualizar seus parâmetros para corrigir as distorções no som durante a reprodução do sinal de áudio, coisa que os outros métodos baseados na aplicação de ruídos e interrupção do sinal não eram capazes de realizar. Os parâmetros para a equalização do ambiente são encontrados em poucas dezenas de segundos e apresentam um grau de convergência apropriado.

Como os trabalhos nessa área estão ainda em estágios iniciais, muitas melhorias e otimizações podem ainda ser alcançadas em trabalhos futuros. É interessante que hajam mais implementações em tempo-real (utilizando processadores dedicados para DSP) com novas técnicas de implementação para se alcançar o melhor que esse método de equalização pode oferecer.

Referências

- [1] Rieta, J.J.; Sanchis, J.M.. **A new method to determine the frequency response of enclosures using masked tones**. Applied Acoustics 67 (2006), págs.: 297–305.
- [2] Rocha, A. F.; Leite, A.; Pinto, F.; Ferreira, A. J. S.. **Adaptive Audio Equalization of Rooms based on a Technique of Transparent Insertion of Acoustic Probe Signals**. Presented at the Audio Engineering Society 120th Convention, Paris, France, 2006 May 20–23.