

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

AVALIAÇÃO DA QUALIDADE SONORA DE COMPRESSORES HERMÉTICOS
UTILIZANDO REDES NEURAS ARTIFICIAIS

DISSERTAÇÃO SUBMETIDA À UNIVERSIDADE FEDERAL DE SANTA
CATARINA PARA OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA
MECÂNICA

CLÁUDIO DE PELLEGRINI

FLORIANÓPOLIS, MARÇO DE 2005

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

AVALIAÇÃO DA QUALIDADE SONORA DE COMPRESSORES HERMÉTICOS
UTILIZANDO REDES NEURAIAS ARTIFICIAIS

CLÁUDIO DE PELLEGRINI

Esta dissertação foi julgada adequada para a obtenção do título de

MESTRE EM ENGENHARIA

ESPECIALIDADE ENGENHARIA MECÂNICA

sendo aprovada em sua forma final.

Arcanjo Lenzi, Ph.D
Orientador

José A. Bellini da Cunha Neto, Dr. Eng.
Coordenador do Curso

Banca Examinadora

Roberto Jordan, Dr. Eng.
Presidente

Mauro Roisenberg, Dr. Eng.

Jorge Cláudio da Silva Pinto, Dr. Eng.

À Cassiane, minha querida esposa,
aos meus pais Cláudio de Marilice
e ao meu irmão Rodrigo.

AGRADECIMENTOS

É com sinceridade que agradeço à minha esposa, aos meus pais e ao irmão pela compreensão, carinho e incentivo ao estudo por todos estes anos.

Externo minha gratidão aos professores, colegas de pesquisa da engenharia mecânica e do programa de pós-graduação em engenharia mecânica da UFSC que diretamente ou indiretamente contribuíram para a realização deste trabalho. Agradeço especialmente aos professores Arcanjo Lenzi e Mauro Roisenberg pelo incentivo e ao amigo Edmar Baars pela motivação e ajuda. Também agradeço à grande contribuição feita por Daniel de Camargo.

Agradeço ainda à Embraco S/A pela disponibilidade de tempo e de equipamentos que permitiram a realização de medições e testes. A eles meus sinceros agradecimentos.

Agradeço também ao suporte financeiro disponibilizado pela CAPES.

SUMÁRIO

AGRADECIMENTOS.....	iii
GLOSSÁRIO.....	vii
LISTA DE FIGURAS E TABELAS	x
RESUMO	xiii
ABSTRACT	xiv
CAPÍTULO 1 - INTRODUÇÃO.....	1
CAPÍTULO 2 - QUALIDADE SONORA.....	6
2.1 - Métricas da Psicoacústica.....	6
2.1.1 - Banda crítica	7
2.1.2 - Curvas de ponderação.....	8
2.1.3 - Loudness	9
2.1.4 - Fluctuation Strength	17
2.1.5 - Roughness.....	18
2.1.6 - Efeito de Mascaramento	20
2.1.7 - Sharpness	23
2.1.8 - Métricas de tom puro.....	25
2.1.9 - Annoyance e Pleasantness	26
2.1.10 - Impulsiveness.....	27
2.1.11 - Outras métricas.....	27
2.2 - Ensaios de júri.....	28
2.2.1 - O ambiente de teste	29
2.2.2 - O júri.....	30
2.2.3 - Preparação da Amostra.....	31
2.2.4 - Preparação do Ensaio	32
2.3 - Métodos de avaliação	33
2.3.1 - Posto de ordem.....	33
2.3.2 - Escala de resposta.....	33
2.3.3 - Comparação pareada	34
2.3.4 - Diferencial semântico	35
2.3.5 - Estimativa de magnitude	35
2.4 - Métodos de análise.....	36
2.4.1 - Estimativa de magnitude, escalas de resposta e diferencial semântico	36
2.4.2 - Comparações pareadas	37

2.4.3 - Ordenamento.....	37
CAPÍTULO 3 - REDES NEURAIS ARTIFICIAIS	39
3.1 - Topologia da rede	40
3.1.1 - Funcionamento dos neurônios	41
3.1.2 - Dinâmica do sistema	41
3.2 - Características das RNAs	42
3.2.1 - Não linearidade	42
3.2.2 - Mapeamento entre a entrada e a saída.....	42
3.2.3 - Adaptatividade	43
3.2.4 - Tratamento de informação contextual.....	43
3.2.5 - Tolerância a falhas	44
3.3 - Modelos de neurônios	44
3.3.1 - O neurônio de McCulloch e Pitts.....	44
3.3.2 - O Perceptron	45
3.3.3 - O modelo do elemento linear adaptativo - Adaline.....	47
3.3.4 - A RNA do tipo perceptron multicamadas ou Multi Layer Perceptron.....	48
3.3.5 - O algoritmo de retropropagação (backpropagation)	49
3.3.6 - O coeficiente de aprendizado α	50
3.3.7 - O problema do mínimo local	51
3.3.8 - Otimização do Perceptron de múltiplas camadas	52
CAPÍTULO 4 - METODOLOGIA DE AVALIAÇÃO SUBJETIVA	54
4.1 - Características gerais do compressor.....	54
4.2 - Sinais de ruído	55
4.3 - Local de teste	56
4.4 - Seleção dos jurados.....	56
4.5 – Ajuste de ganho dos sinais.....	57
4.6 - Apresentação dos sinais.....	58
4.7 - A análise de júri	59
4.8 - Resultados da análise de júri	61
CAPÍTULO 5 - BASE DE DADOS PARA A REDE NEURAL.....	65
5.1 - A elaboração do banco de dados.....	65
5.2 - Dados de classificação subjetiva do ruído.....	66
5.3 - Dados de caracterização objetiva do ruído.....	68
5.4 - A base de dados	71
5.5 - Descrição geral da base de dados.....	72

CAPÍTULO 6 - CONFIGURAÇÃO DA REDE NEURAL ARTIFICIAL	74
6.1 - Ferramenta utilizada.....	74
6.1.1 - Formatação do arquivo de relações e atributos – ARFF	75
6.2 - Parâmetros do sistema.....	76
6.3 - Metodologia de busca da melhor configuração	77
6.4 - Configuração do classificador considerando os atributos dB(A) e loudness	79
6.4.1 - Configuração inicial	79
6.4.2 - Critério de parada do treinamento	80
6.4.3 - Número de neurônios na camada oculta.....	82
6.4.4 - Número de camadas	83
6.4.5 - Momento.....	86
6.4.6 - Taxa de aprendizado.....	87
6.4.7 - Configuração final.....	88
6.5 - Configuração do classificador considerando os atributos dB(A), <i>loudnes</i> , <i>sharpness</i> , <i>roughness e fluctuation strenght</i>	89
6.5.1 - Configuração inicial	89
6.5.2 - Critério de parada do treinamento	89
6.5.3 - Número de neurônios na camada oculta.....	90
6.5.4 - Número de camadas	91
6.5.5 - Momento.....	93
6.5.6 - Taxa de aprendizado.....	93
6.5.7 - Configuração final.....	94
6.5.8 - Avaliação geral das performances das redes neurais artificiais	95
CAPÍTULO 7 - CONCLUSÕES	97
REFERÊNCIAS BIBLIOGRÁFICAS.....	99

GLOSSÁRIO

O objetivo deste é o de esclarecer o significado de alguns termos que definem as métricas da psicoacústica, mas que não possuem uma tradução normalizada nas publicações brasileiras. São mantidos no texto os termos originais em inglês para facilitar o entendimento de outras referências sobre o tema.

Qualidade sonora

<i>“loudness”</i> :	Sonoridade ou audibilidade.
<i>“sharpness”</i> :	Agudeza.
<i>“fluctuation strength”</i> :	Intensidade de flutuação
<i>“roughness”</i> :	Aspereza.
<i>“proeminence ratio”</i> :	Razão de proeminência.
<i>“tone-to-noise ratio”</i> :	Razão de tom para ruído.
<i>“annoyance”</i> :	Incômodo.
<i>“pleasantness”</i> :	Agradabilidade.
<i>“impulsiveness”</i> :	Impulsividade.
High Frequency Factor (HF):	Fator de alta frequência.
Low Frequency Factor (LF):	Fator de baixa frequência.
Spectrum Balance (SB):	Balanco espectral.
Articulation Index (AI):	Índice de articulação.
Speech Interference Level (SIL):	Nível de interferência na fala.
Speech Transmission Index (STI):	Índice de transmissão da fala.

Inteligência artificial

ARFF:	Formatação de arquivos de relações e atributos.
MLP:	<i>Perceptron</i> de múltiplas camadas
RNA:	Redes neurais artificiais.
WEKA:	Ambiente Waikato para análises de conhecimento.
<i>Fold cross validation:</i>	Validação com permuta de conjuntos.
<i>Target:</i>	Resposta desejada.
<i>Feedforward:</i>	Propagação adiante.
<i>Hard limiting:</i>	Degrau unitário.
<i>Backpropagation:</i>	Retropropagação.
<i>Gradient descent:</i>	Gradiente descendente.
<i>LMS - Least Mean Square:</i>	Erro médio quadrático
ADALINE:	Elemento Linear Adaptativo
<i>Flat plateau:</i>	Mínimo local próximo do início do treinamento.
$e(t)$	Erro associado ao neurônio
α	Taxa de aprendizado
$o(t)$	Saída associada ao neurônio
i	Entrada do neurônio
β	Inclinação da função sigmóide
θ	Valor de limiar (<i>Threshold</i>)
w_i	Peso sináptico associado ao neurônio i
Δw	Valor de correção aplicada ao peso sináptico

$\partial f(nf)$	Derivada da função de ativação
$\delta(t)$	Derivada da função de erro
α	Coefficiente de aprendizado
m	Coefficiente de momento

LISTA DE FIGURAS E TABELAS

Figuras

Figura 2.1 – Largura de banda crítica em função da frequência. [6].....	7
Figura 2.2 – Curva formada pela relação entre escala de banda crítica e frequência. [6].....	8
Figura 2.3 – Curvas de Ponderação A, B e C. [7]	9
Figura 2.4 – Curvas de igual sonoridade para tons puros em campo livre. [6].....	10
Figura 2.5 – Atenuação, a_d , necessária para produzir a mesma sonoridade para um tom puro em campo difuso e em campo livre.[5].....	11
Figura 2.6 – Exemplo de gráfico utilizado para determinação do “loudness” pelo método de Zwicker. [8]	14
Figura 2.7 – Gráfico para a determinação dos valores de índice de loudness para o cálculo de “loudness” do pelo método de Stevens. [8].....	16
Figura 2.8 – “fluctuation strength” de um ruído de banda larga com amplitude modulada com 60 dB de nível de pressão sonora e 40 dB de intensidade de modulação. [6].....	18
Figura 2.9 – Nível de um tom de teste mascarado por um ruído branco de um dado nível de densidade em função da frequência do tom de teste. A curva com contorno inferior representa o limiar da audição. [6].....	21
Figura 2.10 – Nível do tom de teste mascarado por um ruído de banda crítica com nível de 60 dB e frequências centrais de 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz e 8 kHz. [6].....	22
Figura 2.11 – Nível do tom de teste mascarado por um ruído de banda crítica com frequência central de 1 kHz em diferentes níveis. [6].....	23
Figura 2.12 – Representação do conceito do “Tone-to-noise ratio”, sendo W_t o nível de pressão sonora do tom puro e W_n o nível de pressão sonora da banda com largura de banda crítica com a ausência do tom puro.[6]	25
Figura 2.13 – Representação do conceito do “Proeminence ratio”, sendo W_c a potência da banda crítica em torno do tom puro e W_i e W_u as potências das bandas críticas adjacentes. [6]	26
Figura 3.1 – Exemplo de modelo de uma RNA.	40
Figura 3.2 – Exemplo de redes simples de McCulloch-Pitts. O exemplo a) ilustra a lógica replicadora, b) a lógica E e c) a lógica OU. [18].	45
Figura 3.3 – Separação linear de dois padrões no espaço, por uma rede de dois perceptrons.	47

Figura 3.4 – Diagrama de um perceptron de múltiplas camadas de três camadas clássico.	49
Figura 4.1 – Microfone binaural utilizado no processo de ajuste.....	58
Figura 4.2 – Modelo de tabela apresentada aos jurados para seleção dos ruídos apresentados.	61
Figura 4.3 – Distribuição das avaliações subjetivas para os sinais considerados mais agradáveis.	62
Figura 4.4 – Distribuição das avaliações subjetivas para os sinais considerados semelhantes.	63
Figura 4.5 – Distribuição das avaliações subjetivas para os sinais considerados mais incômodos.....	63
Figura 5.1 – Fluxograma de elaboração do banco de dados.	66
Figura 5.2 – Sinal de ruído no domínio do tempo.	68
Figura 5.3 – Espectro de um terço de oitava utilizado como dado de entrada para a rede neural.	69
Figura 5.4 – Valores de Loudness utilizados como dado de entrada para a rede neural.	70
Figura 5.5 – Valores de fluctuation strength utilizados como dado de entrada para a rede neural.	70
Figura 5.6 – Valores de roughness utilizados como dado de entrada para a rede neural.	71
Figura 5.7 – Formato da tabela utilizada como banco de dados para a rede neural artificial. .	72
Figura 6.1 – Exemplo de formatação ARFF.	75
Figura 6.2 – Avaliação do número de épocas de treinamento.....	80
Figura 6.3 – Avaliação do percentual do conjunto de validação a ser utilizado para interromper o treinamento.	81
Figura 6.4 – Avaliação do número de neurônios na camada oculta.	82
Figura 6.5 – Avaliação do número de neurônios camada oculta próximo ao ponto ótimo.....	83
Figura 6.6 – Avaliação do número de camada desbalanceadas em 36 neurônios.	84
Figura 6.7 – Avaliação do número de camada balanceadas em 36 neurônios.	85
Figura 6.8 – Avaliação do parâmetro momento.	86
Figura 6.9 – Avaliação do parâmetro taxa de aprendizado.	87
Figura 6.10 – Avaliação do número de neurônios na camada oculta.	90
Figura 6.11 – Avaliação da tendência do erro médio para a variável validation treshhold igual a um.....	91
Figura 6.12 – Avaliação do número camadas desbalanceadas em 120 neurônios.	92
Figura 6.13 – Avaliação do número camadas balanceadas em 120 neurônios.	92
Figura 6.14 – Avaliação do parâmetro de momento.....	93
Figura 6.15 – Avaliação do parâmetro taxa de aprendizado.	94

Tabelas

Tabela 5.1 – Tabela de distribuição de votos para os sons considerados mais agradáveis.....	67
Tabela 5.2 – Tabela descritiva da base de dados.....	73
Tabela 6.1 – Erro médio das configurações de RNAs finais.....	95

RESUMO

A necessidade crescente por pesquisas e trabalhos de desenvolvimento relacionados à redução do ruído de produtos tem demonstrado que os níveis de ruído medidos com o uso da curva de ponderação A, em dB (A), não representam de forma satisfatória a resposta subjetiva das pessoas ao ruído. Métricas de qualidade sonora são usadas para a caracterização psicoacústica do ruído e proporcionam melhores avaliações pelo fato de considerarem informações mais detalhadas nos domínios do tempo e da frequência. Este trabalho apresenta uma análise da qualidade sonora de compressores herméticos. Para este tipo de produto foi observada boa correlação entre as respostas do júri, níveis em dB (A), e resultados de *loudness*. Os resultados do júri indicaram maior sensibilidade de percepção a detalhes da composição do ruído, quando comparados às métricas de qualidade sonora. Entretanto, o uso de um júri representa um procedimento mais elaborado, de maior custo e tempo de análise. Para reduzir estas desvantagens uma rede neural artificial foi desenvolvida e testada com o objetivo de simular o comportamento do júri a estas classes de produto. Os resultados apresentam uma concordância razoável aos obtidos por um júri real, sendo incentivada a continuidade deste procedimento.

ABSTRACT

The increasing needs for research and development works on products noise reduction has lead to observations that A-weighted sound pressure and sound power measurements do not represent customers subjective response to noise in a complete way. Sound quality metrics were developed for the psychoacoustic characterization of sound as it is perceived and provide an improved assessment since it considers more detailed contents in time and frequency domains. This work presents a sound quality analysis of noise generated by hermetic compressors. It was observed a close relationship between jury response, A-weighted noise levels, and loudness results obtained for this product. Jury results showed higher sensitivity to specific details of noise composition compared to current sound quality metrics. Jury analysis, however, is a more elaborated and time consuming procedure. In order to reduce such disadvantages an artificial neural network was developed and tested with the objective of representing the jury response to this class of product. Results are in reasonably good agreement with those obtained by a real jury. This has encouraged continuing the development of this technique.

CAPÍTULO 1 - INTRODUÇÃO

O desenvolvimento tecnológico de produtos industrializados, juntamente com os crescentes níveis de exigência dos consumidores finais vêm, ao longo destes anos, abrindo espaço para características, até então, pouco valorizadas. Dentre estas características vem se destacando o ruído.

A implantação de normas para legalização dos níveis de ruído em ambientes foi um grande passo para que se voltassem as atenções para esta característica peculiar, que também tem se tornado um importante fator para uma vida saudável. Recentemente alguns eletrodomésticos tiveram os níveis de ruído incorporados em suas especificações técnicas. Estes níveis, no entanto, indicam apenas a potência sonora, não sendo suficiente para indicar o quão incômodo ou agradável é aquele ruído, isto é, apesar de indicar a energia sonora, não sugere nenhuma informação quanto à sua qualidade sonora.

A qualidade sonora de um produto é um conceito que agrupa e representa vários outros conceitos, tornando assim a sua definição uma tarefa difícil, levando em conta as muitas áreas de aplicação. A definição mais comumente aceita pelos estudiosos da área de qualidade sonora é a seguinte “Qualidade sonora é a reação perceptiva ao som de um produto que reflete a reação do ouvinte de quanto satisfatório o som do produto é: o mais satisfatório, a melhor qualidade sonora” [1].

Cada produto apresenta características e atribuições que são traduzidas pelos seus sons, e estes são caracterizados pela sua distribuição espectral e temporal resultando na percepção final pelos consumidores, sendo que as empresas têm como objetivo atuar sobre estes ruídos de forma a torná-los mais aceitáveis [2]. No caso de sistemas de refrigeração, nos quais a principal fonte de ruído é o compressor, a redução do incômodo sonoro produzido por este produto está relacionada à melhora da qualidade sonora do ruído gerado pelo compressor.

A questão agora passa a ser como se pode avaliar a qualidade sonora destes produtos.

A maneira mais efetiva é a apresentação de um grupo de sons a um conjunto de ouvintes e solicitar que os mesmos façam julgamentos em relação aos sons, porém com o cuidado de realizar um experimento que forneça resultados com validade estatística e conceitual. Esta forma é a análise de júri.

Ensaio com júri, no entanto, envolve tempo e custos consideráveis para sua realização, levando-se em conta a seleção e os treinamentos que devem ser aplicados ao júri, de forma que possa realizar um julgamento de maior isenção.

Outra forma de avaliação do incômodo causado, ou mesmo a otimização de suas características, se baseia na aplicação de conceitos da psicoacústica, aliados à análise de júri. A psicoacústica é um ramo da psicologia experimental que estuda e identifica as correlações entre os estímulos físicos do som e as reações, ou percepções, das pessoas ao som.

Um exemplo de aplicação da análise de júri para o desenvolvimento da psicoacústica é o julgamento do “*loudness*” de tons. A percepção do “*loudness*” é um dos grandes referenciais na psicoacústica, tanto que a escala física do “*loudness*” é utilizada quase que universalmente para prever como os ouvintes irão julgar o “*loudness*” dos produtos.

Outro parâmetro utilizado, e também conhecido, é o mascaramento, que é o encobrimento de um som sobre o outro. Uma importante aplicação do mascaramento é na interferência da fala, podendo-se prever com boa precisão o quanto a inteligibilidade da fala será prejudicada pelo ruído de fundo, seja em uma biblioteca ou em um praça de alimentação, baseando-se num teste de júri de mascaramento e percepção da fala.

Visando aproximar a percepção humana do ruído e, assim, estudá-lo e otimizá-lo, utiliza-se de um conceito físico chamado de curva de ponderação A, ou curva dB (A), que por sua vez aproxima os valores de ruído medidos com a forma na qual o sistema auditivo responde. Porém, a curva de ponderação A considera apenas a energia do sinal, não informando características como transientes e mascaramentos no domínio tempo, e ainda modulações no sinal, surgindo daí a necessidade de ferramentas que supram tais deficiências.

O trabalho realizado por Nunes [2], mostrou claramente esta deficiência. O experimento baseou-se na utilização de quatro sistemas de refrigeração com os mesmos níveis de ruído dB (A), mas que apresentaram diferenças quando avaliados por um júri, tornando clara a limitação da curva dB (A) em relação às nuances específicas dos ruídos.

Neste trabalho foram analisadas as métricas disponíveis em um software comercial de análise de qualidade sonora (Sound Quality v.3.6, fabricado pela empresa Brüel & Kjaer®), aplicadas a sinais de ruído de compressores de refrigeração. Concluiu-se que as métricas não apresentam boa correlação com a avaliação dos jurados. Por exemplo, as métricas de tom puro, “*tone-to-noise*” e “*proeminence ratio*”, não apresentam qualquer correlação para todas as configurações analisadas, havendo grandes variações em seus valores. Em suma, a determinação de correlações entre as métricas da psicoacústica e os resultados do júri nas avaliações dos ruídos dos compressores de refrigeração apontou a existência de correlação apenas para os valores de “*loudness*” e dB (A), embora as amplitudes dos valores dB (A) e do “*loudness*” apresentem variações diferentes.

Tendo em vista os resultados apresentados neste trabalho, é válido ressaltar que formas alternativas de avaliação da qualidade sonora de compressores de refrigeração, bem como ferramentas auxiliares, devem ser buscadas e desenvolvidas.

O processo de qualificação sonora, realizado por um conjunto de jurados, pode ser resumido como um processo de classificação, que tem como critério percepções subjetivas. Para tal tarefa, entenda-se a seleção, uma ferramenta de origem relativamente recente que vem sendo utilizada com grande sucesso, é a Rede Neural Artificial ou RNA, bastando para tanto a utilização adequada dos parâmetros envolvidos no processo de sua utilização.

Assim, este trabalho tem como objetivo principal o desenvolvimento e a utilização de uma rede neural artificial para a qualificação sonora de compressores herméticos de refrigeração, baseando-se em uma análise de júri preliminar.

Numa primeira etapa o trabalho é concentrado na escolha, coleta e edição de sinais de ruído. A seqüência envolve o estudo e a escolha do método de aplicação dos sinais a um júri, a realização de uma análise de júri propriamente dita e a análise dos resultados obtidos. A partir de então, o foco passa a ser o tratamento destas informações para a formação de um banco de dados, que será usado como entrada na rede neural artificial. A etapa final envolve basicamente a escolha e o desenvolvimento das redes neurais, seguidas de uma análise de sensibilidade dos parâmetros envolvidos, e finalizando com a avaliação dos resultados. Isto para que, num segundo momento, se faça desnecessária a avaliação destes ruídos por uma análise de júri. Abrindo-se assim, a possibilidade do uso desta metodologia para o desenvolvimento da qualidade sonora de novos produtos ainda em fase de concepção.

O Capítulo 2 desta obra apresenta as métricas da psicoacústica mais utilizadas atualmente, com suas curvas e definições, os cuidados na preparação e execução dos ensaios e as possíveis formas de avaliação dos resultados.

No Capítulo 3 é apresentada uma revisão teórica e bibliográfica das redes neurais, descrevendo seus conceitos, suas definições, suas classificações, suas formas de treinamento e suas aplicações.

Os procedimentos utilizados para a aplicação do conceito de qualidade sonora estão descritos no Capítulo 4. A forma de aquisição e de tratamento dos dados, a metodologia de avaliação subjetiva e os resultados da aplicação dos conceitos de qualidade sonora para compressores também são apresentados neste capítulo.

A etapa de elaboração da base de dados a ser utilizada para o treinamento da rede neural é apresentada no Capítulo 5. O Capítulo 6 descreve o desenvolvimento da rede neural artificial (RNA), abordando a avaliação da melhor configuração de rede, análises de sensibilidade da rede, a avaliação dos atributos mais significativos e por fim a configuração final da rede proposta.

No Capítulo 7 estão as conclusões do trabalho e algumas sugestões para trabalhos futuros.

CAPÍTULO 2 - QUALIDADE SONORA

Saber porque o som da água que pinga da torneira irrita, enquanto que o barulho de milhões de gotas de chuva aumenta o prazer de adormecer, é uma questão que está na base da psicoacústica, uma ciência relativamente antiga que associa a psicologia com a percepção auditiva, em busca do que se chama qualidade sonora. [3].

Existem registros de realização de testes de qualidade sonora, na indústria automobilística, que datam do início do século passado. Buzinas de automóveis eram instaladas em painéis e acionadas diante de um júri que escolhia aquelas com melhor sonoridade. Outro teste realizado, também com o auxílio de jurados, baseava-se na avaliação do ruído da passagem de automóveis, sendo este um indicativo de preocupação com a poluição sonora. [3]. O teste com júri é típico da psicoacústica e deve ser orientado por critérios psicológicos. Isto porque, um mesmo ruído pode soar de uma forma se o jurado estiver tranqüilo e em ambiente agradável, e de outra se a pessoa estiver estressada. É necessário avaliar se os juízes estão em situação de neutralidade.

2.1 - Métricas da Psicoacústica

A área de qualidade sonora utiliza-se de conceitos de um ramo da acústica dedicada ao estudo das correlações entre as características físicas do som e as reações subjetivas ocasionadas pelo mesmo, a psicoacústica. [4].

Com o intuito de contornar o problema, que envolve a avaliação do som através do júri, foram criados conjuntos de fórmulas e tabelas denominados métricas objetivas. Estas possuem a função de completar os ensaios com júri e mensurar as correlações existentes.

2.1.1 - Banda crítica

A base para o desenvolvimento de parâmetros psicoacústicos é o conceito de banda crítica, que é definida como uma banda com valores de largura de banda com 100 Hz, para frequências abaixo de 500Hz e de aproximadamente $(0,2 \cdot f)$ para as demais frequências. Esta se baseia no fato de que o sistema auditivo humano analisa espectros largos em segmentos que correspondem às bandas críticas. A Figura 2.1 apresenta as larguras de banda crítica em função da frequência. As aproximações para as faixas de baixas e altas frequências são indicadas pelas linhas pontilhadas. [5]

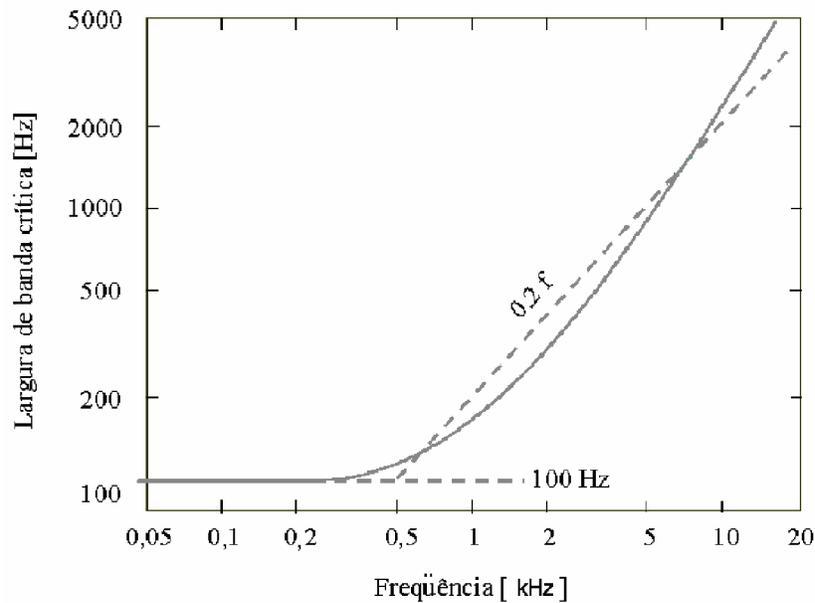


Figura 2.1 – Largura de banda crítica em função da frequência. [6]

Uma vez definidos os valores de banda crítica podem-se definir os valores da escala de taxa de banda crítica, e assim, adicionando uma banda crítica à próxima de forma que o limite superior da banda crítica anterior corresponda ao limite inferior da próxima. A escala resultante da união destes pontos é chamada de taxa de banda crítica, que vai de 0 a 24 e possui a unidade “Bark”, em memória ao cientista Barkhausen [5]. A relação entre a taxa de

banda crítica e a frequência é importante para o entendimento de várias características do ouvido humano (Figura 2.2).

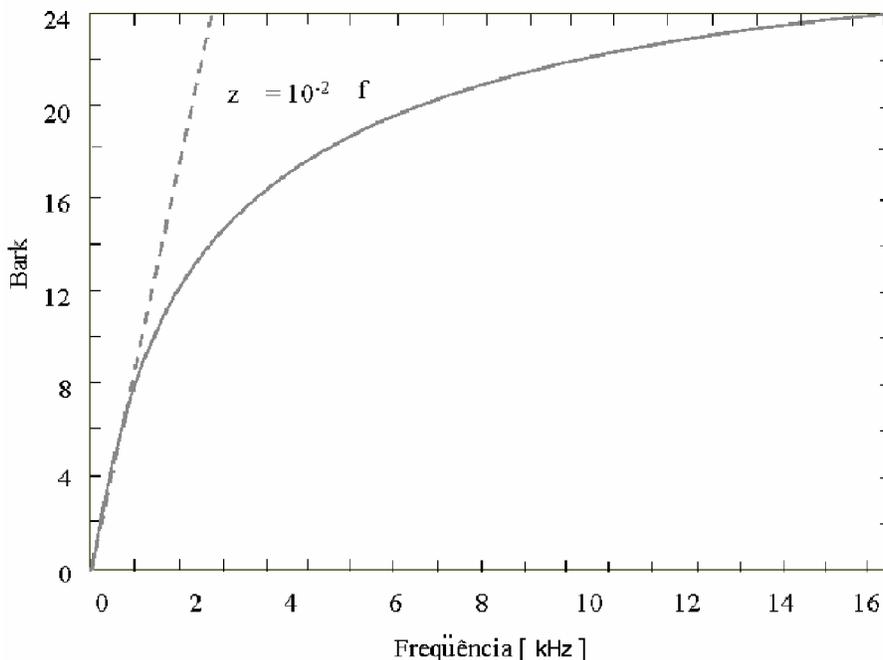


Figura 2.2 – Curva formada pela relação entre escala de banda crítica e frequência. [6]

2.1.2 - Curvas de ponderação

Os espectros de ruídos, físicos, conforme medidos por um instrumento, são diferentes do espectro subjetivo, ouvido por nós. Os efeitos da orelha humana são considerados, nos instrumentos de medição (medidores de nível de pressão sonora), através de três curvas de ponderação, A, B e C.

As curvas de ponderação não são parâmetros psicoacústicos, no entanto, são formas comumente adotadas para a representação de ruídos. Dentre estas, a curva de ponderação A é a mais popular, pois é utilizada diretamente e serve de base para o cálculo de índices usados para expressar vários tipos de ruído, como por exemplo, ruído ocupacional, ruído rodoviário, ruído aeroportuário, conforto interno de ambientes, entre outros.

A curva de ponderação A é determinada utilizando-se como referência um tom padrão de 1 kHz com nível de pressão sonora de 40 dB. A partir deste ponto a frequência é variada e são determinados os valores de níveis de ruído que proporcionam a mesma sonoridade. Da mesma forma podem ser obtidas as curvas de ponderação B e C, que diferem entre apenas pelo nível de referência, 70 dB para a curva de ponderação B e 100 dB para a curva de ponderação C [7].

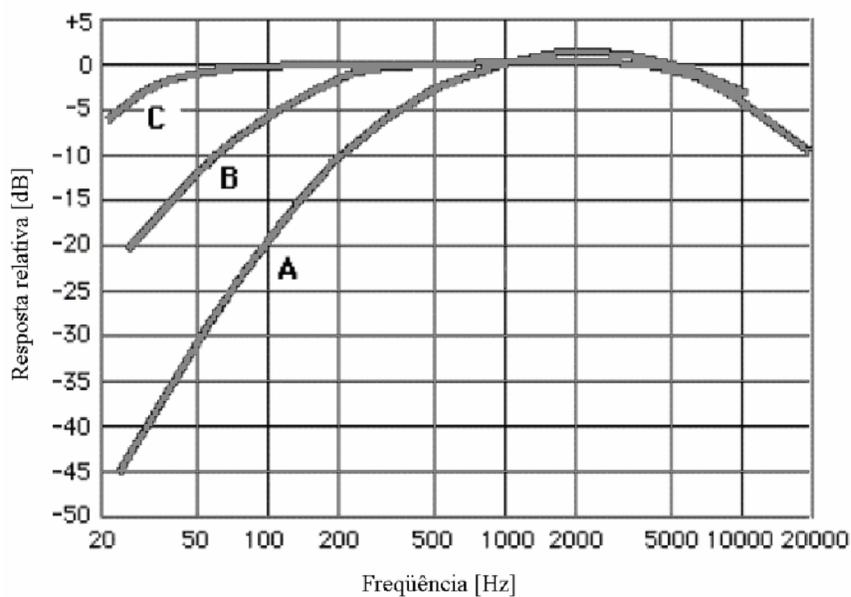


Figura 2.3 – Curvas de Ponderação A, B e C. [7]

2.1.3 - Loudness

Trata-se da métrica mais utilizada. O termo “loudness” pode ser traduzido como sonoridade, e refere-se à percepção da intensidade do som.

O nível de “loudness” de um som é o nível de pressão sonora igual a 40 dB de um tom puro de 1 kHz de uma onda plana com incidência normal ao ouvido que seja tão sonoro quanto o som que se quer avaliar, e sua unidade é o “phon”. Quando se quer expressar o valor de “loudness” a unidade é o “sone”. A escala “sone” é linearmente relacionada à sensação

auditiva, tornando-a de fácil interpretação. Um som duas vezes mais alto que outro possui o dobro do valor de “sones” original ou, de outra forma, possui um valor 10 “phons” maior.

A união dos pontos de mesmo “loudness”, na região de audição, é chamada de linha de igual sonoridade. Linhas de igual sonoridade para tons puros são mostradas na figura abaixo, e em função de sua definição, todas as curvas devem possuir valores tendendo ao nível da pressão sonora em 1 kHz. Assim a curva de igual sonoridade para 40 “phons” deve tender para um valor de 40 dB em 1 kHz. Seguindo ainda esta definição tem-se a curva do limiar da audição, onde o limite de sensação do “loudness” é alcançado, com valor de 3 “phons”, já que o limite do silêncio corresponde a 3 dB em 1 kHz. [5].

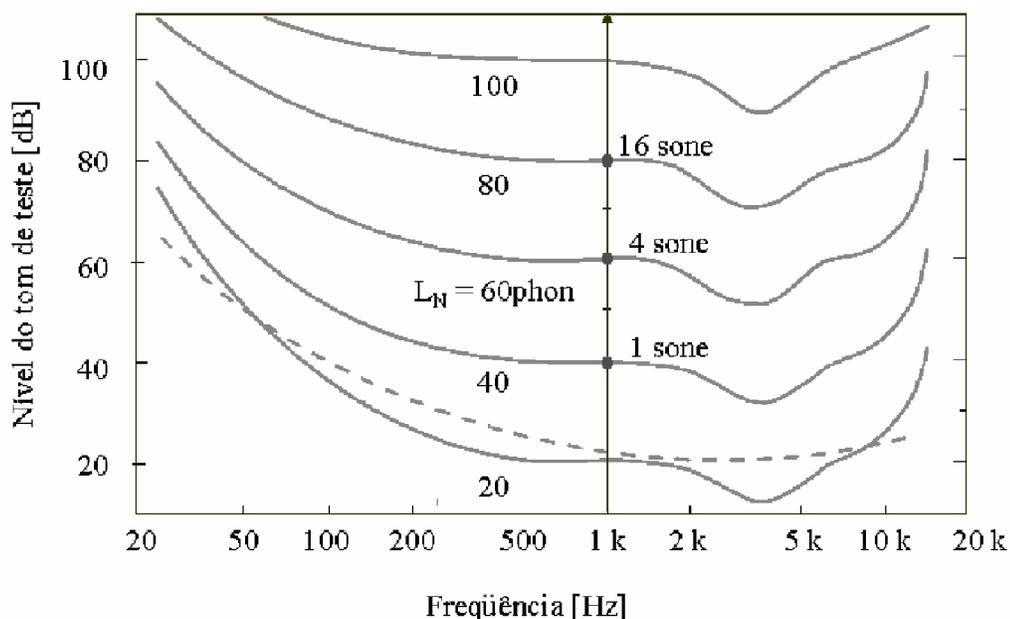


Figura 2.4 – Curvas de igual sonoridade para tons puros em campo livre. [6]

Os valores do gráfico anterior foram obtidos para tons puros em campo livre, isto é, para ondas planas de incidência normal. No entanto, em muitos casos, o campo sonoro pode ser difuso, no qual o som incide em todas as direções. O sistema auditivo não é igualmente sensível ao som vindo de diferentes direções e a dependência da direção também está relacionada com a frequência. Por esta razão, as curvas de igual sonoridade para campos

sonoros normais e para campos sonoros difusos são diferentes. Pode-se expressar esta diferença por um fator de atenuação, denominado fator a_d . A aplicação deste fator a um ruído em campo difuso faz com este produza o mesmo “loudness” deste ruído em campo normal. A Figura 2.5 mostra a relação desta atenuação com a frequência. Observa-se que nas baixas frequências a atenuação é desprezível, pois o sistema auditivo funciona como receptor onidirecional.

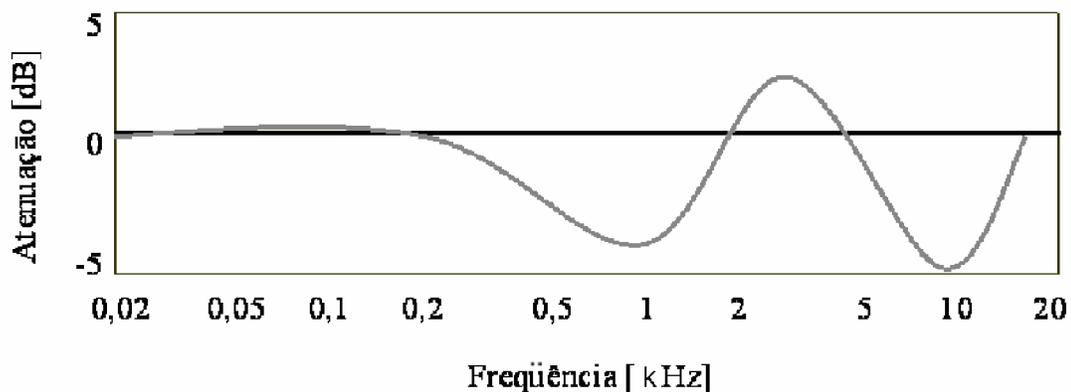


Figura 2.5 – Atenuação, a_d , necessária para produzir a mesma sonoridade para um tom puro em campo difuso e em campo livre.[5]

Assim, em 1 kHz, onde a atenuação é de -3 dB, tem-se que o nível de pressão sonora de um tom de 1 kHz num campo difuso deve ser 3 dB menor do que a pressão sonora de um de 1 kHz em campo normal, para produzir a mesma sensação de sonoridade. Em relação às altas frequências, a atenuação, a_d , incrementa até alcançar o limite de 2 dB a 2,5 kHz.

Os resultados, a_d , apresentados acima, foram obtidos com medições de tons harmônicos, com ruídos de banda estreita. Portanto, a frequência apresentada como abscissa na Figura 2.5 deve ser atribuída como frequência central do ruído de banda estreita. Utilizando os dados apresentados na Figura 2.5 juntamente com os apresentados na Figura 2.4, as linhas de igual sonoridade para o campo difuso podem ser facilmente construídas.

As curvas de igual sonoridade vistas até aqui, tanto para tons puros quanto para ruídos de banda, apresentam somente interesse acadêmico. Fletcher e Munson [2] propuseram um procedimento para o cálculo a partir de medidas físicas de sons complexos formados por um determinado número de tons. O seu método, entretanto, não é utilizado em função da sua complexidade e pelo fato de a grande maioria dos sons de interesse prático tenderem a espectros de banda larga e não de tons puros.

Gates [2] e depois Beranek [2] propuseram que uma simples soma do “*loudness*” em sonos em bandas de oitava deveria fornecer, com uma aproximação razoável, o “*loudness*” percebido por sons complexos compostos por mais de uma oitava de um ruído aleatório. Foi assumido que um ruído aleatório em bandas de oitava tendo o mesmo nível global de pressão sonora que o de um tom puro com mesma frequência central, possuiriam mesmo “*loudness*”.

Porém estes métodos não se mostraram operacionais ou com valores não totalmente confiáveis. Assim, para a determinação de valores de “*loudness*” atualmente se utilizam dois métodos principais: o Método de Stevens e o Método de Zwicker, sendo o primeiro utilizado normalmente para bandas de oitava e o segundo para bandas de terço de oitava. Ambos os métodos são normalizados e os procedimentos encontram-se na norma ISO 532 [8].

Um terceiro método, o de Kryter [9], tem sido empregado para estimar o ruído externo de aeronaves (ruído de aeroportos), enquanto que os de Zwicker e Stevens tem sido utilizados para ruídos de uma forma geral da fala e música.

Método de Zwicker

Estudos realizados por Fletcher e Munson sugeriram que o “*loudness*” é proporcional ao número de impulsos nervosos que chegam por segundo ao cérebro pelos nervos auditivos. Perceberam, ainda, que quando dois tons eram impostos simultaneamente acontecia a soma do “*loudness*” dos tons, sendo que isto ocorria para todo um grupo de tons numa certa banda

de freqüência. A largura deste grupo de freqüência foi levantada por Fletcher e Munson, porém os valores determinados por Zwicker, com testes subjetivos de mascaramento e “loudness” revelou valores mais exatos, que são os valores de largura de banda crítica.

A definição de banda crítica como sendo aproximadamente iguais às bandas de terço de oitava para freqüências acima de 280 Hz, e para bandas menores, como igual aos grupos de bandas de terço de oitava é considerada na norma ISO 532 [8].

Baseando-se no conjunto de gráficos desenvolvidos por Zwicker, o procedimento de cálculo é dividido nas seguintes etapas: [10].

- i) Seleção do gráfico em função do tipo de campo e valor máximo do nível de banda de terço de oitava medido, e da marcação dos valores de nível de pressão sonora em cada gráfico, sendo que cada um conte as curvas na forma de degraus que relacionam os valores dos níveis de cada banda de terço de oitava (L_t).

Os níveis medidos devem ser marcados nos gráficos para as freqüências de banda acima de 280 Hz na forma de linhas horizontais, sendo que as freqüências de corte de cada banda de terço de oitava correspondem às abscissas e os níveis indicados nos degraus.

Como nas baixas freqüências as bandas críticas são mais largas que as bandas de terço de oitava, é feito o agrupamento de bandas de baixas freqüência da seguinte forma: para freqüências com freqüência de corte até 90 Hz, tem-se uma banda crítica L_1 ; para bandas com freqüência entre 90 Hz e 180 Hz tem-se L_2 , e para as duas bandas entre 180 Hz e 280 Hz tem-se L_3 .

A forma de combinação das freqüências é apresentada da seguinte forma:

$$L_2 = 10 \cdot \log_{10} \left(\frac{\text{anti log } L_{100}}{10} + \frac{\text{anti log } L_{125}}{10} + \frac{\text{anti log } L_{160}}{10} \right), \quad (2.1)$$

onde L_{100} é o nível de pressão sonora medido na banda de terço de oitava na frequência central 100 Hz, L_{125} para 125 Hz e L_{160} para 160 Hz.

- ii) O segundo passo indica como fazer a conexão dos pontos marcados. Se os degraus formados pelas linhas horizontais estiverem crescendo juntamente com a frequência, as linhas horizontais adjacentes deverão ser ligadas por linhas verticais entre suas abscissas comuns. Quando o nível na próxima frequência superior for menor, a curva para baixo é desenhada com uma inclinação descendente interpolada entre as linhas tracejadas do gráfico, iniciando do fim da linha horizontal da direita.
- iii) Este passo indica que a área determinada deve ser transformada num retângulo de mesma área, tendo uma base com valor igual ao comprimento do gráfico. Assim, a altura do retângulo indica o nível de “loudness”, em “phons”, ou em sones, a partir das escalas adjacentes ao gráfico.

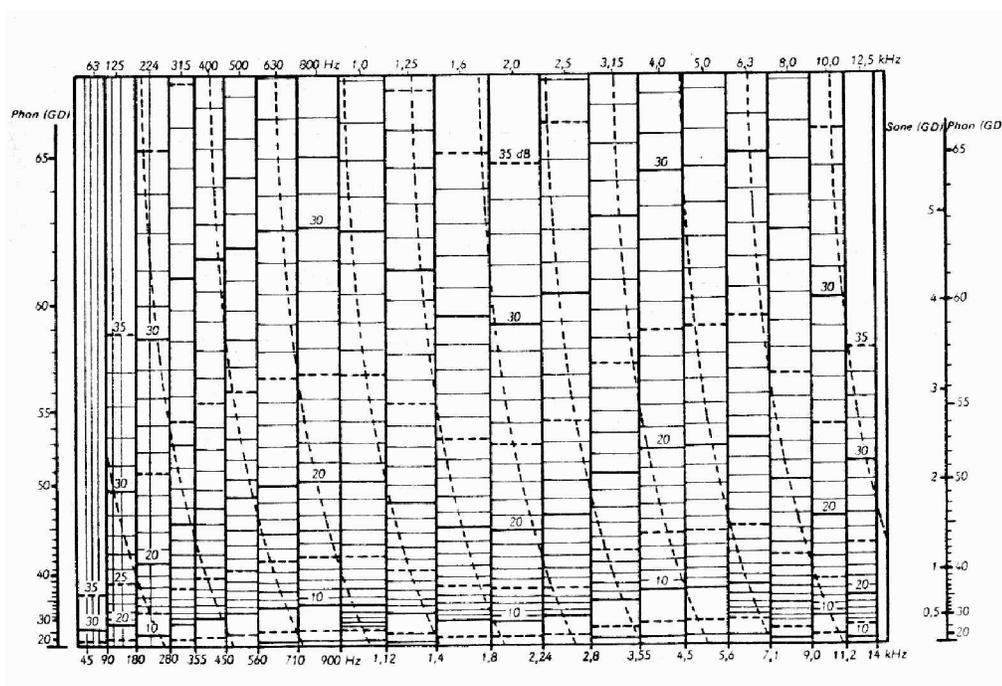


Figura 2.6 – Exemplo de gráfico utilizado para determinação do “loudness” pelo método de Zwicker.

Método de Stevens

Utilizando-se das curvas de igual sonoridade já existentes para a determinação do *loudness* de ruídos aleatórios em bandas de oitava, Stevens desenvolveu novos procedimentos para a determinação do *loudness* total de espectros contínuos de sons.

O método de Stevens demonstra ser mais preciso na predição e no julgamento do *loudness* de sons complexos compostos de ruídos aleatórios do que o método de simples adição dos valores de sone das bandas individuais. [10].

A fórmula geral de Stevens é dada pela soma do valor, em sone, da banda com maior *loudness* com uma fração proporcional da soma do valor em sone das bandas restantes:

$$loudness = s_m + f \cdot \left(\sum s - s_m \right), \quad (2.2)$$

sendo: s_m o maior valor, em sone, apresentado pelas bandas de frequência,

f , a fração proporcional que depende da largura de banda,

s , o valor, em sone, das bandas restantes.

Stevens determinou frações para bandas de oitava ($f = 0,3$), meia oitava ($f = 0,2$) e terço de oitava ($f = 0,15$).

O Método Mark IV, adotado pela ISO 532 [9] como procedimento a ser utilizado para o cálculo do *loudness* em ruídos de oitava, é baseado no método descrito acima com uma pequena modificação realizada por Stevens. Os procedimentos e a fórmula para o cálculo neste novo método são idênticos ao método anterior, sendo os valores individuais de cada banda encontrados em um gráfico que indica as linhas índice de *loudness*, que substituem os valores em sones. Estas curvas são semelhantes às de igual sonoridade.

Para a determinação do nível de *loudness*, aplica-se o valor de S_r na fórmula:

$$P = 40 + 10 \log_2(S_r), \quad (2.3)$$

A Figura 2.7 apresenta o gráfico para a determinação dos valores de *loudness index* para o cálculo do “*loudness*” pelo método de Stevens

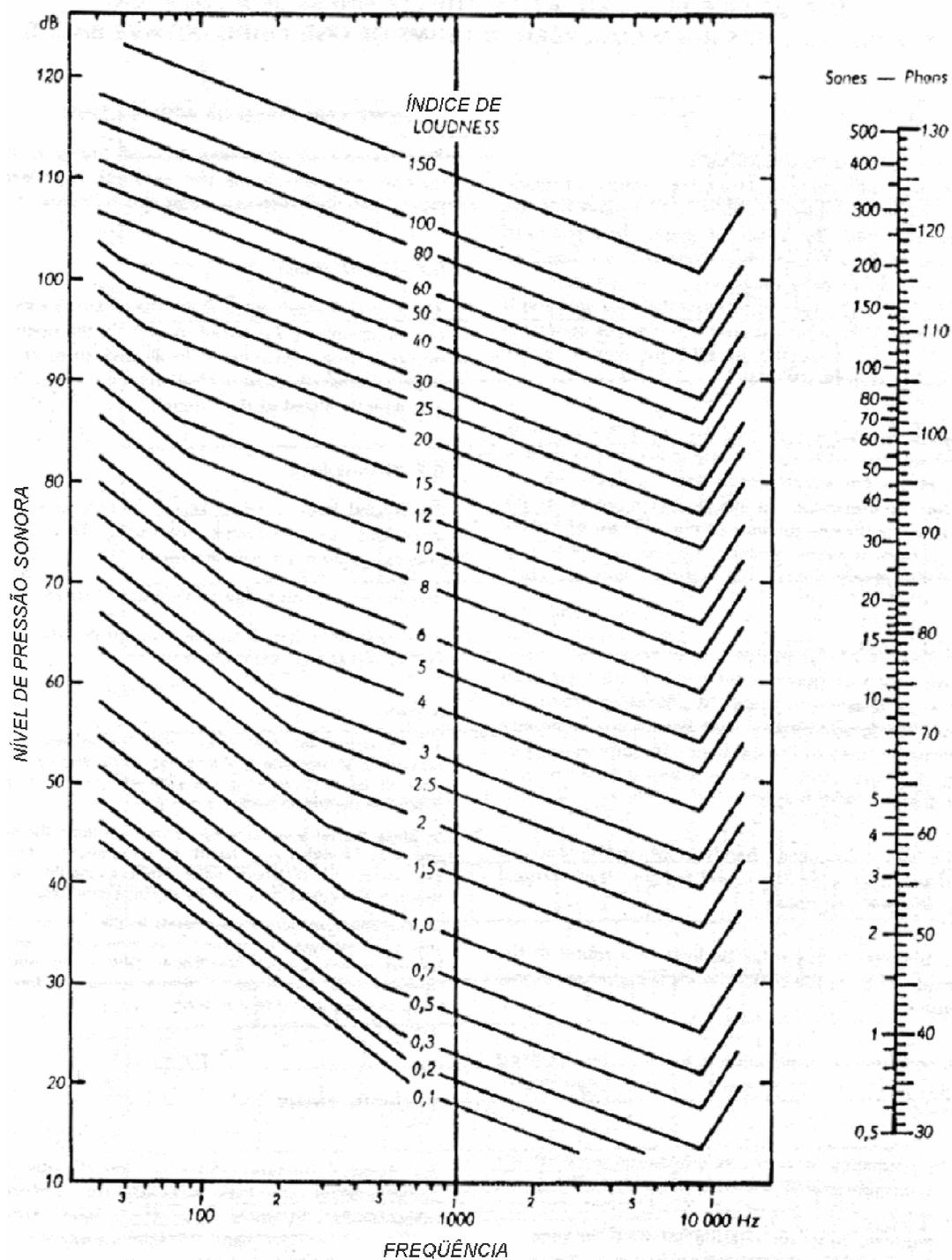


Figura 2.7 – Gráfico para a determinação dos valores de índice de *loudness* para o cálculo de “*loudness*” do pelo método de Stevens. [8]

2.1.4 - Fluctuation Strength

A “*fluctuation strength*” representa o grau de flutuação para sensações auditivas de baixa frequência de modulação, até cerca de 16 Hz. Para frequências de alta frequência de modulação, a sensação “*roughness*” é percebida. Para frequências de modulação em torno de 20 Hz, há uma transição suave entre a sensação de audição do “*fluctuation strength*” e do “*roughness*”, que apresenta uma grande região de indefinição entre estas sensações.

Um tom modulado na faixa 1 a 16 Hz. causa uma falsa sensação de variação da amplitude ao longo do tempo, sendo percebido mais alto do que um som contínuo com a mesma amplitude. A unidade utilizada para representar a “*fluctuation strength*” é o “*vacil*”, sendo que as frequências de modulação em torno de 4 Hz são mais sensíveis ao ouvido. Frequências de modulação acima e abaixo desse valor são menos sensíveis, sendo esse o motivo pelo qual as sirenes possuem frequências de modulação próximas dessa faixa. O modelo de fluctuation strength proposto por Zwicker e Fastl é um dos mais utilizados [11]:

$$FS = 0,008 \cdot \frac{\int_0^{24Bark} \Delta L \cdot dz}{\left(\frac{f_{mod}}{f_0} + \frac{f_0}{f_{mod}} \right)}, \quad (2.4)$$

onde: FS é a fluctuation strength em vacil;

ΔL , a amplitude de modulação;

f_{mod} , a frequência de modulação;

f_0 , igual a 4 Hz.

A Figura 2.8 apresenta a relação do “*fluctuation strength*” com a frequência de modulação, sendo este normalizado ao máximo valor do som.

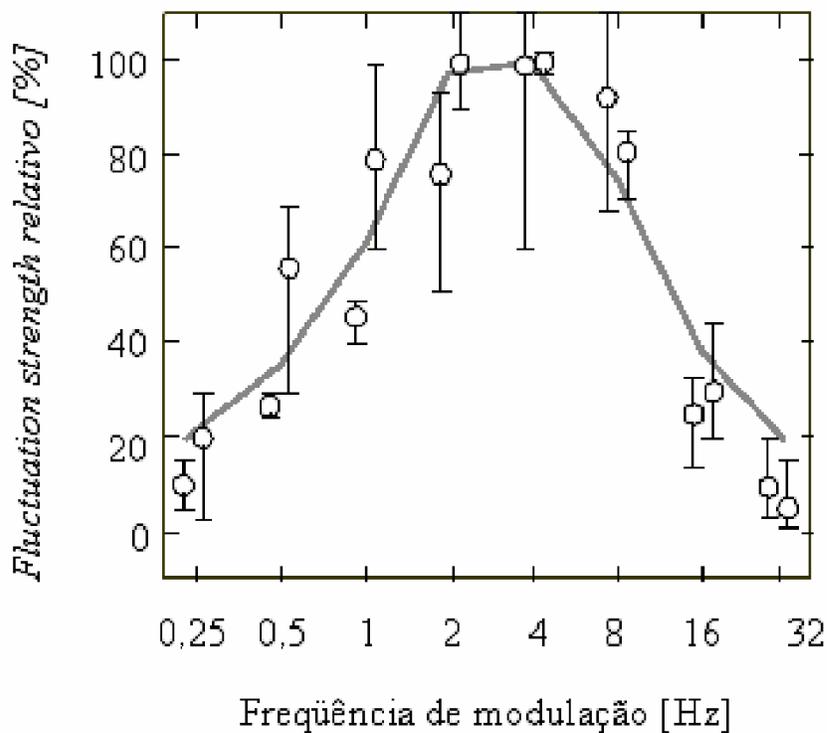


Figura 2.8 – “*fluctuation strength*” de um ruído de banda larga com amplitude modulada com 60 dB de nível de pressão sonora e 40 dB de intensidade de modulação. [6]

A justificativa para o máximo valor de “*fluctuation strength*” estar em 4 Hz se baseia na variação do envelope temporal da fala fluente, pois a taxa normal da fala, de 4 sílabas por segundo, é usualmente produzida, levando-a a uma variação do envelope temporal numa frequência de 4 Hz. Também justifica o porquê da excelente correlação entre a fala e o sistema auditivo.

2.1.5 - *Roughness*

Assim como a métrica anterior, o “*roughness*” também está relacionado com a modulação do sinal. Utilizando um tom de 1kHz com amplitude 100% modulada e variando o valor da frequência iniciando das baixas para as altas frequências, podem-se identificar três áreas de sensações distintas. A primeira região é a do “*fluctuation strength*”, que alcança seu

valor máximo em 4 Hz e decresce até em torno de 16 Hz, onde a sensação do “*roughness*” começa a se tornar perceptível, alcançando o seu valor máximo em frequências de modulação próximas a 70 Hz e decrescendo para maiores valores de frequência.

Com o objetivo de quantificar o “*roughness*”, um valor de referência foi definido. Partindo-se do termo em latim “*asper*”, que caracteriza o que se define de “aspereza sonora”, definiu-se para o “*roughness*” de 1 asper, um tom puro com 60 dB de intensidade em 1 kHz, modulado 100 % numa frequência de modulação de 70 Hz. Três parâmetros são importantes para a determinação do “*roughness*”: a amplitude de modulação, o grau de modulação e a frequência de modulação

A sensação de “*roughness*” se deve a uma mudança relativamente rápida produzida pela frequência de modulação na região compreendida entre 15 Hz e 300 Hz, não havendo necessidade de uma modulação periódica exata, mas apenas que o espectro de modulação esteja entre 15 Hz e 300 Hz. E é por esta razão, que a maioria dos ruídos de banda estreita com um som dito “áspero” não apresentam uma variação periódica no envelope ou na frequência. O “*roughness*”, assim como o “*sharpness*” e o “*fluctuation strength*” é também uma sensação que pode ser considerada individualmente, ignorando as demais sensações. [12].

O *roughness* pode ser calculado através da seguinte equação:

$$R = 0,3 \cdot f_{\text{mod}} \cdot \int_0^{24 \text{Bark}} \Delta L_{\text{MASK}}(z) \cdot dz, \quad (2.5)$$

onde: R é a *roughness* expresso em asper;

ΔL_{MASK} , a amplitude de modulação levando em consideração os efeitos de mascaramento;

f_{mod} , a frequência de modulação.

2.1.6 - Efeito de Mascaramento

O efeito de mascaramento é definido como sendo a sobreposição de um sinal sobre o outro, tendo como consequência o encobrimento, total ou parcial, da percepção subjetiva do primeiro. Pode-se verificar este efeito em peças musicais, no momento em que um instrumento inicia a sua parte sobrepondo-se à melodia executada por outro, podendo desta forma estar mascarando o primeiro.

O ouvido humano também apresenta este efeito de mascaramento, pois muitos ruídos são mascarados em função de que a membrana basilar no ouvido interno, ao ser excitada por um ruído, não responde em apenas um ponto localizado, mas sim numa área, afetando a percepção de ruídos que se encontrem nesta faixa.

Com o intuito de qualificar este efeito deve-se determinar o limiar do mascaramento, sendo, para isto, utilizado um tom puro, chamado de tom de teste, e determinando-se o nível de pressão sonora necessário para que se torne audível quando comparado com um sinal mascarador. Também é possível que o mascaramento ocorra na forma de um sinal mascarador impulsivo, que é emitido antes do início do som a ser mascarado, ou como um sinal que é desligado durante a emissão do sinal a ser mascarado, gerando uma sensação de redução do mesmo. Estes efeitos são denominados, respectivamente, de pré e pós mascaramento.

Para estudos acústicos utilizam-se ruídos com características conhecidas para determinação de valores padrões, como o ruído branco, que possui uma densidade espectral constante com a frequência, e o ruído rosa, que apresenta densidade espectral inversamente proporcional à frequência. Para a determinação das características de mascaramento também são utilizados estes ruídos.

Os limites de mascaramento de um tom de teste mascarado por um ruído branco em vários níveis de densidade, que é o valor da intensidade sonora numa banda de 1 Hz, são apresentados em linhas contínuas na Figura 2.9. A linha cheia com contorno inferior representa o limiar da audição.

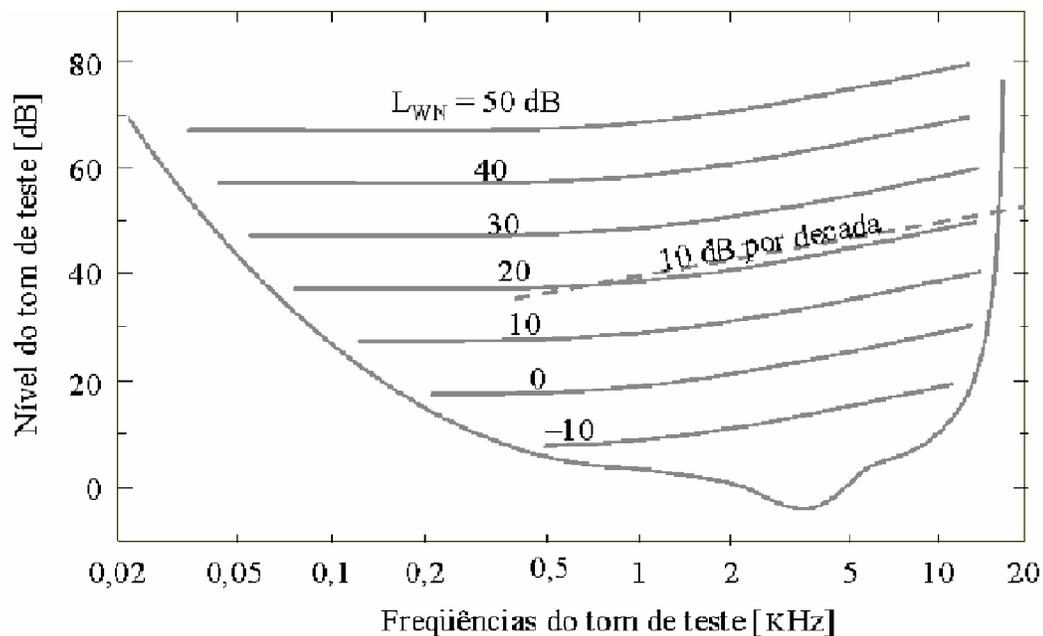


Figura 2.9 – Nível de um tom de teste mascarado por um ruído branco de um dado nível de densidade em função da frequência do tom de teste. A curva com contorno inferior representa o limiar da audição. [6]

Embora o ruído branco possua uma densidade espectral independente da frequência, percebe-se que acima das frequências em torno de 500 Hz as curvas de mascaramento crescem em função do incremento da frequência, numa taxa de 10 dB por década. Percebe-se também que para os extremos de frequência as curvas de mascaramento acabam se confundindo com o limiar de audição.

A banda estreita, aqui aplicada, supõe ruídos com largura de banda iguais ou menores do que as larguras de banda crítica, ou seja, com cerca de 100 Hz abaixo de 500 Hz e de 20% da frequência central para frequências acima de 500 Hz.

Tons puros mascarados por ruídos com larguras de bandas críticas com frequências centrais de 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz e 8 kHz, são mostrados na Figura 2.10.

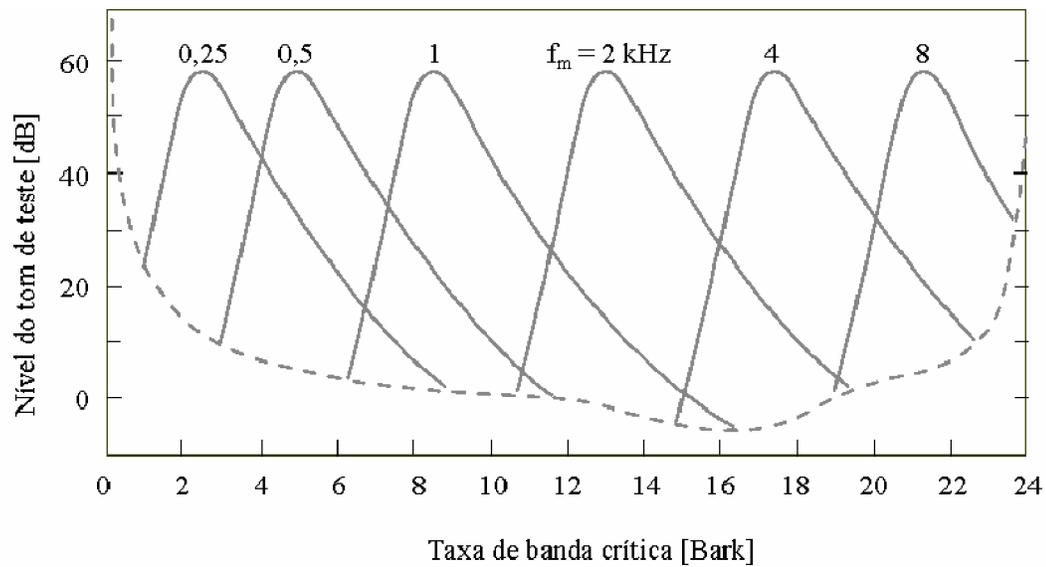


Figura 2.10 – Nível do tom de teste mascarado por um ruído de banda crítica com nível de 60 dB e frequências centrais de 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz e 8 kHz. [6]

Na Figura 2.11 é apresentada a relação do limite de mascaramento do nível de um ruído centrado em 1 kHz. Percebe-se que todos os limites de mascaramento mostram uma ascendência muito rápida das baixas para as altas frequências antes do mascaramento máximo ser alcançado. A curva deste aumento é independente do nível do ruído mascarador, e o máximo sempre se encontra 3 dB abaixo do nível de ruído de mascaramento.

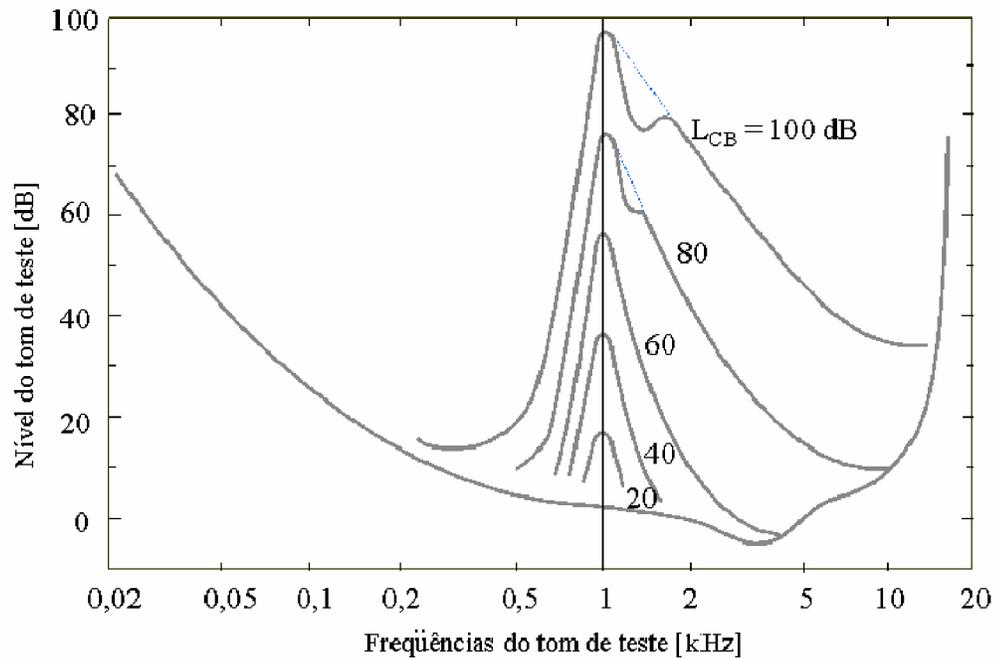


Figura 2.11 – Nível do tom de teste mascarado por um ruído de banda crítica com frequência central de 1 kHz em diferentes níveis. [6]

A frequência do limite mascarado é dependente do nível, ou seja, afastando-se do valor máximo, o limite de mascaramento se reduz rapidamente para baixos valores, isso para baixos e médios níveis mascaradores. Entretanto, para altos níveis mascaradores, a curva em direção às altas frequências torna-se cada vez menos acentuada. Os declives indicados na Figura 11 para níveis mascaradores de 80 dB a 100 dB se devem aos efeitos não lineares no sistema auditivo, que cria a diferença de audibilidade de ruídos criados pela interação entre o tom de teste e o ruído de banda estreita.

2.1.7 - Sharpness

Esta métrica é considerada um parâmetro importante na avaliação de timbre e sua unidade é representada pelo “*acum*”, cujo significado, do latim, é agudo. Sons com uma

grande quantidade de componentes em alta frequência possuem um alto valor de “*sharpness*”.

A sensação quantificada pelo “*sharpness*” pode ser considerada separadamente das demais, ou seja, pode-se comparar o “*sharpness*” de um ruído diretamente com o de um outro ruído, sendo que o principal fator que irá diferenciar os seus valores é a variação do seu envelope espectral. Assim, ruídos que apresentam espectros contínuos podem apresentar o mesmo valor de “*sharpness*” de ruídos que apresentam espectro com muitas variações. [11].

A dependência do “*sharpness*” com nível pressão sonora pode ser ignorada como primeira aproximação, especialmente se a diferença entre níveis não for muito grande. Isto se justifica pelo fato de que, para um incremento de 30 dB para 90 dB, o “*sharpness*” incrementa por um fator de dois. Também é pequeno o efeito da relação de largura de banda, tanto com bandas pequenas quanto com a escala de banda crítica, já que nenhuma diferença no “*sharpness*” pode ser determinada se um tom ou cinco tons for utilizado para gerar uma banda crítica, ou mesmo quando é utilizado um ruído com banda crítica para a comparação.

Assim, os parâmetros de maior importância na influência do “*sharpness*” são o conteúdo espectral e a frequência central de sons de banda estreita. De maneira a fornecer valores quantitativos, um ponto de referência e uma unidade foram definidos. Um “*acum*” foi definido como um ruído de banda estreita com largura de banda crítica com frequência central em 1 kHz e com um nível de 60 dB.

Atualmente existe uma série de modelos utilizados para o cálculo de “*sharpness*”, sendo que destes, o mais comum e aceito é o modelo proposto por Von Bismarck [12]:

$$S = 0,11 \cdot \frac{\int_0^{24Bark} z \cdot L'(z) \cdot g(z) \cdot dz}{\int_0^{24Bark} L(z) \cdot dz}, \quad (2.6)$$

onde: S representa o sharpness em acum;

z representa as bandas críticas “Frequenzgruppe”

$L'(z)$ representa o “loudness” específico para cada banda crítica;

$g(z)$ representa um fator de escala correspondente a 1 para as bandas críticas de 1 a 16 Bark e 4 para a banda crítica de 24 Bark. Entre 16 e 24 Bark, tal fator cresce de 1 a 4

2.1.8 - Métricas de tom puro

Estas métricas foram desenvolvidas para suprir a necessidade da indústria de informática em analisar as componentes de tons puros emitidos por seus equipamentos. São elas: “*tone-to-noise ratio*” e “*proeminence ratio*”

O “*tone-to-noise ratio*” pode ser definida como a relação entre a potência do tom e a potência da banda crítica centrada no tom.

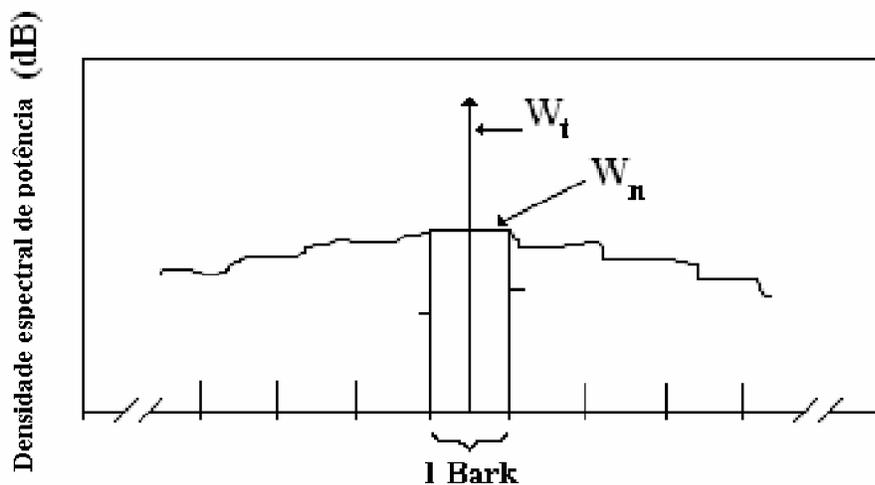


Figura 2.12 – Representação do conceito do “*Tone-to-noise ratio*”, sendo W_t o nível de pressão sonora do tom puro e W_n o nível de pressão sonora da banda com largura de banda crítica com a ausência do tom puro.[6]

O “*proeminence ratio*” é definida como sendo a relação entre potência do tom e a potência média das bandas críticas adjacentes.

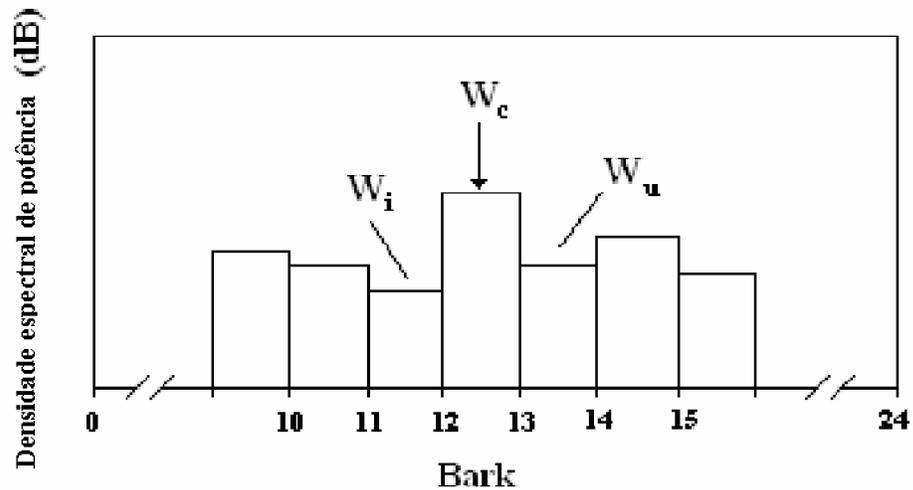


Figura 2.13 – Representação do conceito do “Proeminence ratio”, sendo W_c a potência da banda crítica em torno do tom puro e W_i e W_u as potências das bandas críticas adjacentes. [6]

A avaliação destes parâmetros é feita relacionando o quanto “proeminente” é o sinal, sendo que para o “proeminence ratio” diferenças acima de 7 dB são consideradas proeminentes, e para o “tone-to-noise ratio” diferenças acima de 6 dB já são consideradas proeminentes, ou seja, nestes casos as componentes de tom puro afetam na avaliação dos jurados.

2.1.9 - Annoyance e Pleasantness

Com o objetivo de aliar as características das métricas acima mencionadas em uma única medida global foram desenvolvidas as métricas denominadas “Annoyance” e “Pleasantness”. Normalmente estas métricas são resultados de combinações de métricas tradicionais.

A “annoyance” é uma métrica desenvolvida por Zwicker, esta leva em consideração o “sharpness”, a “fluctuation strength”, o “loudness” e o horário do dia em que é feita a

análise (dia ou noite), sendo o “*loudness*” a métrica mais importante para o cálculo. A “*annoyance*” é dada em termos de unidades de “*annoyance*”, au, [13].

De forma semelhante, o “*pleasantness*” é uma métrica que leva em consideração o “*sharpness*”, “*roughness*”, “*loudness*”..

2.1.10 - *Impulsiveness*

O ouvido humano possui a característica de mascarar alguns sons e ruídos a quais está submetido de acordo com as características temporais dos mesmos.

Os músculos situados no interior do ouvido possuem a capacidade de enrijecerem-se quando os ruídos que o atingem possuem níveis muito elevados. [10]. Esse enrijecimento provoca uma diminuição na amplitude percebida do ruído; no entanto, existe um tempo para que tais efeitos se manifestem. No caso de ruídos impulsivos, os músculos do sistema auditivo ficam sem “tempo de reação” e a maneira como percebemos o ruído é completamente diferente. Uma das métricas de qualidade sonora mais conhecidas que tenta quantificar esse efeito, é o “*impulsiveness*”.

2.1.11 - *Outras métricas*

São comuns os casos em que as métricas apresentadas não representam de forma satisfatória um fenômeno sonoro com a exatidão desejada. Assim, engenheiros de ruído geralmente recorrem às métricas específicas para os diferentes produtos. Tais métricas são obtidas através da combinação de outras métricas e índices, baseando-se em métodos empíricos e utilizando grandes quantidades de dados gerados através de avaliações subjetivas por sessões de júri, [10]. Entre alguns dos índices mais utilizados destacam-se:

1. Fator de crista e Kurtose – utilizados para ruídos impulsivos;

2. High Frequency Factor (HF), Low Frequency Factor (LF) e Spectrum Balance (SB);
3. High Frequency Level (HFL) – avalia o conteúdo espectral acima de 1 kHz;
4. Métricas de inteligibilidade de fala – em várias aplicações a qualidade sonora de um produto está associada diretamente com a interferência provocada na fala. Existe uma série de parâmetros que podem representar esse fator, tais como: índice de articulação (articulation index – AI), Speech Interference Level (SIL), Preferred Speech Interferation Level (PSIL), Speech Transmission Index (STI), Rapid Speech Transmission Index (RASTI);
5. Harshness, rumble, muddiness – utilizados quando se deseja representar algumas componentes tonais do espectro;
6. Gradiente harmônico – alguns parâmetros acústicos relacionados ao sentimento de “esportividade” e “prazer acústico” podem ser interessantes para o desenvolvimento de métricas para determinados tipos de produtos. Estudos indicaram que parâmetros mensuráveis relacionados ao gradiente harmônico de componentes tonais podem ser associados a esses sentimentos.

2.2 - Ensaio de júri

Quando se objetiva realizar uma qualificação sonora de algum produto, o ensaio com júri é a primeira e mais importante etapa deste processo. Contudo, tal ensaio exige uma série de preparativos e condições especiais que têm como objetivo principal minimizar a influência dos parâmetros que não estejam relacionados à subjetividade do ensaio.

2.2.1 - O ambiente de teste

O primeiro requisito para o ambiente de teste é o baixo valor do nível de ruído de fundo, para que o mesmo não interfira no julgamento dos jurados. Assim, deve-se tomar um cuidado especial com possíveis fontes de ruído, tais como computadores, lâmpadas fluorescentes e outros, para que não afetem os resultados dos testes.

Quanto à apresentação, às características da sala de teste, é recomendado que esta seja tão natural quanto possível, sendo equipada com cadeiras e fones de ouvido confortáveis, tendo cores neutras e iluminação moderada, com o objetivo de evitar ao máximo que o jurado desvie sua atenção e acabe por afetar o resultado dos testes. Também deve-se levar em conta a temperatura da sala, estando em níveis considerados agradáveis, entre 20 e 24 °C e com umidade relativa em torno de 50%. [10].

Para ambientes de teste reverberantes, devem-se tomar alguns cuidados com relação à reprodução do som, pois a utilização de alto-falantes em qualquer sala que não uma câmara anecóica, fará com que a resposta em frequência desta fique sobreposta à resposta em frequência dos alto-falantes, fazendo assim com que o som reproduzido perca a fidelidade ao som originalmente gravado. No caso de realizar testes com mais de um jurado simultaneamente, a situação tende a piorar, uma vez que a percepção do som para cada um será diferente. Este conjunto de fatores leva à conclusão de que a reprodução através de fones de ouvido seja a melhor opção.

Muitos autores da área defendem a importância de estímulos visuais da sala de teste e que toda melhoria realizada, com o objetivo de aproximar o ambiente de teste a um ambiente real, é válida. Baseando-se nessa idéia, a utilização de simuladores tem ganho espaço nos últimos anos. A intenção é criar um ambiente virtual com o produto ainda na fase de projeto. Assim, podem-se aplicar testes de avaliação subjetiva no simulador e, se necessário, realizar possíveis modificações no produto antes da obtenção do protótipo, diminuindo custos e aumentando a chance de se ter um produto boa qualidade sonora. Levando em consideração o

fato de o contexto ser importante na percepção auditiva, jurados normais (entenda-se não peritos) necessitam experimentar as sensações provocadas pelo ruído e pela vibração de uma forma tão próxima quanto possível da realidade.

2.2.2 - O júri

É desejável que o jurado represente com alguma fidelidade o consumidor final. Assim, a escolha do jurado deve basear-se na sua familiaridade com o produto, na sua experiência em ensaios e em alguns aspectos como idade, sexo, e status econômico.

Um importante fator a ser considerado é a familiaridade com o produto. Parte dos julgamentos pode ser influenciada pelas suas expectativas em relação ao produto. Com o objetivo de evitar esse inconveniente, os jurados devem ser escolhidos de tal forma que suas expectativas não comprometam os resultados dos ensaios. É comum as empresas utilizarem seus funcionários, uma vez que estes geralmente possuem um conhecimento sobre todos os produtos da linha de produção.

Com relação à experiência com os ensaios, pode-se afirmar que sujeitos mais experientes conseguem perceber com maior clareza algumas nuances dos sons apresentados, enquanto que sujeitos inexperientes apresentam alguma dificuldade. Por outro lado, peritos no assunto tendem a valorizar aspectos que podem não ser importante ao usuário final. Uma vez que, para a escuta da grande maioria dos sons, não é necessário um elevado grau de experiência, jurados com pouca ou nenhuma experiência podem ser utilizados.

Outros fatores, como idade, sexo, e status econômico, também devem ser considerados. Por exemplo, ao analisar-se um veículo de luxo, deve-se levar em consideração que a maioria dos consumidores finais possuem uma faixa etária elevada e são, predominantemente, homens. No caso de um eletrodoméstico, a amostragem deve ser feita predominantemente com mulheres de média idade.

Por fim deve-se levar em conta o número de avaliadores, pois este está diretamente relacionado à representatividade dos resultados. Um número adequado de jurados para ensaios realizados com funcionários é de 25 a 50, no entanto são necessários 75 a 100 jurados no caso de utilizarem-se potenciais consumidores, [10]. É muito importante que os jurados se familiarizem aos procedimentos, sons apresentados e ao ambiente de teste. É desejável que todos os sons sejam apresentados ainda na fase de treinamento, realizada antes do julgamento subjetivo propriamente dito. Para tarefas mais complexas, um treinamento mais extensivo torna-se necessário.

2.2.3 - Preparação da Amostra

Levando em consideração a complexidade do aparelho auditivo humano, muitas nuances podem representar a diferença de qualidade sonora entre dois produtos distintos. Gravações e ou reproduções mal executadas na amostra do som podem arruinar completamente um ensaio.

Na aquisição dos dados, recomenda-se a utilização da técnica de gravação binaural com faixa dinâmica maior que 90 dB, sendo que a gravação das amostras deve ser realizada durante a utilização do produto e em ambientes com baixo valor de nível de ruído de fundo. Os sons que não fazem parte da análise devem ser eliminados da amostra através de edição ou exclusão dos dados contaminados. Na necessidade de avaliarem-se outros aspectos psicoacústicos não relacionados à amplitude do ruído, faz-se necessária a modificação do “*loudness*”. Isso é feito ajustando a amplitude das amostras até que todas possuam o mesmo valor de “*loudness*”. Essa técnica é particularmente interessante quando a estratégia de ensaio é a de comparação pareada.

No que se refere ao tipo de equalização, existem basicamente três tipos: *free-field* (FF), *diffuse-field* (DF), *independent of direction* (ID). Levando em consideração que, na

maioria dos casos, os campos sonoros analisados não são nem livres e nem difusos, a equalização ID torna-se a opção mais correta das três citadas. A maioria das cabeças artificiais existentes no mercado possui os tipos de equalização mencionados.

2.2.4 - Preparação do Ensaio

Nesta etapa, alguns cuidados especiais devem ser tomados, como por exemplo a ordem de apresentação dos sons. O descuido deste ponto pode introduzir erros sistemáticos ao processo de avaliação e levar a resultados errôneos. Alguns métodos de ensaio, como o das comparações pareadas, já levam em consideração este fator, minimizando sua influência na análise dos dados.

Outro fator importante é a duração do ensaio, que deve ser considerada na medida em que o júri passe a se sentir desconfortável. O tempo aconselhável para a duração máxima dos ensaios é de até 40 minutos. [10]

Como explicado anteriormente, os fones de ouvido constituem uma maneira mais efetiva de apresentar o mesmo estímulo a cada júri, uma vez que podem ser calibrados e equalizados individualmente para que as suas respostas se tornem equivalentes.

Um outro ponto a ser comentado diz respeito às apresentações visuais apresentadas aos avaliadores. Para resultados mais autênticos os sons devem ser reproduzidos no local em que foram gravados ou em ambientes de simulação que se aproxime o máximo possível do ambiente real. Qualquer alteração feita com o objetivo de incrementar o contexto visual no qual o som é reproduzido é considerada como melhoria. Na impossibilidade da reprodução do som no ambiente real ou em simuladores, estímulos visuais, como vídeos demonstrando o funcionamento dos produtos, ou até uma figura, podem controlar a expectativa dos jurados, mantendo o foco destes no produto sob teste.

2.3 - Métodos de avaliação

O objetivo da avaliação subjetiva através de júri é fazer com que os resultados sejam representativos das opiniões dos consumidores ou usuários finais de um produto. Alguns dos métodos para avaliação subjetiva são descritos agora.

2.3.1 - *Posto de ordem*

A aplicação deste método é bastante simples. Solicita-se aos jurados que ordenem os sons de acordo com algum critério de avaliação. No entanto, o grau de dificuldade desta tarefa cresce consideravelmente à medida que o número de sons aumenta. Por isso, o tamanho da amostra é usualmente baixo, por volta de cinco sons.

O fato desse método não dar nenhuma informação quantitativa sobre os sons, revela uma de suas desvantagens. Assim, em uma análise é possível concluir, por exemplo, que o som A é menos incômodo que o som B, contudo não é possível quantificar essa diferença. Em virtude disso, a correlação com parâmetros objetivos não pode ser realizada, fazendo com que tal método sirva apenas para análises rápidas ou para determinar a preferência do consumidor.

2.3.2 - *Escala de resposta*

A avaliação dos jurados é feita atribuindo uma nota para cada um dos sons reproduzidos. Geralmente os sons são reproduzidos em uma determinada seqüência e não são feitas repetições. É um método rápido e simples. Além disso, fornece uma informação quantitativa direta sobre as diferenças entre um som e outro. Um ponto a ser considerado é a sua deficiência no que diz respeito à padronização das escalas de resposta, sendo que é de se esperar que os jurados apresentem réguas distintas de avaliação. Vale ainda lembrar que não

existe nenhuma razão para crer que as escalas de resposta possuem correlação direta com as métricas objetivas, uma vez que as escalas podem variar de um jurado para outro.

2.3.3 - Comparação pareada

O método baseia-se na apresentação dos sons em pares, sendo que os jurados devem realizar julgamentos relativos, baseando-se no par apresentado. Existem basicamente três tipos de avaliações baseadas neste método.

A avaliação de detecção determina que o jurado deva escolher qual dos sons no par contém o sinal a ser detectado. É um método utilizado para determinação de limiares. Por exemplo deseja-se saber até que limite um tom puro é mascarado por um ruído de banda larga [10].

No método de avaliação, os avaliadores realizam um julgamento subjetivo escolhendo um dos sons apresentados no par segundo algum critério de avaliação. Quando forem utilizados critérios de avaliação mais específicos, como por o exemplo o “*sharpness*”, devem ser tomados alguns cuidados para garantir que o jurado compreenda o significado dos termos. O procedimento de escolha é repetido até que todos os pares tenham sido avaliados, sendo que a replicação dos testes também é praticada.

A terceira metodologia de avaliação, diferentemente das duas avaliações mencionadas, solicita que os avaliadores estimem a similaridade entres os sons do mesmo par através de uma escala. Todos os pares possíveis são avaliados dessa maneira. Depois de todas as avaliações, uma grade numerada é colocada sobre a linha e as marcas feitas pelos jurados são convertidas em números.

A principal desvantagem dos métodos de comparação pareada é que o número de pares cresce com o quadrado do número de sons, tornando o processo de avaliação demorado e custoso. Nesses casos, é possível fazer uma análise com blocos incompletos, porém isto

exige uma certa experiência do pesquisador uma vez que resultados satisfatórios só podem ser obtidos com uma seleção apropriada dos pares que formarão os blocos.

2.3.4 - Diferencial semântico

O método das comparações pareadas é focado, unicamente, em um atributo dos sons. No Diferencial Semântico essa avaliação pode ser realizada com múltiplos atributos, usualmente oito a doze, de uma só vez.

Tal avaliação é realizada através de uma escala, contendo de cinco a nove pontos e possuindo um par de adjetivos bipolares. O jurado escolhe um dos pontos da escala de acordo com a que melhor representa sua impressão. A escolha dos pares semânticos é uma tarefa importante, necessitando que elas sejam apropriadas para a aplicação. Revistas, jornais e anúncios são boas fontes para se coletarem os pares, porém a maioria dos psicoacústicos iniciou a procura de adjetivos através de descrições verbais espontâneas.

2.3.5 - Estimativa de magnitude

Neste método o jurado deve atribuir um número a algum atributo do som. Teoricamente não existe limite para a faixa de números a ser avaliado. Essa falta de limite constitui uma vantagem sobre os métodos de comparação pareada e diferencial semântico uma vez que todos os valores atribuídos estão dentro da escala. Por outro lado, esta vantagem também acaba acarretando em problema, uma vez que não existe uma padronização nas respostas dos jurados e, assim, estes podem ter respostas bastante diferentes. Para jurados inexperientes, é um método de difícil execução, necessitando treino e prática. Sendo este, portanto, mais indicado para avaliadores experientes.

2.4 - Métodos de análise

2.4.1 - Estimativa de magnitude, escalas de resposta e diferencial semântico

Os três métodos de avaliação acima descritos, de estimativa de magnitude, de escalas de resposta e de diferencial semântico, fornecem informações importantes do ponto de vista estatístico, uma vez que possibilitam a comparação das respostas entre os jurados.

O método de estimativa de magnitude, por envolver a criação de uma escala própria de cada jurado, necessita que se faça uma normalização antes de qualquer análise estatística. A análise da distribuição deve levar em consideração os seguintes fatores:

- i) medida da localização – medida utilizada para quantificar a tendência central de uma distribuição. Basicamente existem três medidas da localização: média, mediana e modo;
- ii) medida da variabilidade – utilizada para descrever o espalhamento dos dados da distribuição. Essa medida pode ser estimada através do alcance, ou seja, a diferença entre o menor e maior valor, da variância e do desvio padrão;
- iii) medida da forma – utilizada para quantificar a natureza das distribuições resultantes dos dados.
- iv) teste de normalidade – muitos testes estatísticos assumem a normalidade dos dados, assim é desejável que se tenha conhecimento se esta hipótese é ou não preenchida. Os testes de Kolmogorov-Smirnov (K-S tests) são comumente usados para checar quão bem uma distribuição teórica pode representar uma distribuição experimental. A enorme quantidade de dados gerados em ensaios através de júri dificulta a análise. Nesses casos, técnicas gráficas podem tornar esse processo menos oneroso.

2.4.2 - *Comparações pareadas*

Neste método, costuma-se avaliar apenas a repetitividade do jurado, indicador da percentagem de todas as comparações que possuem a mesma resposta na primeira e na segunda verificação, e a consistência de suas respostas, indicador do quão bem os julgamentos dos pares podem ser representados em construções de ordem maiores.

Realizando ensaios com o conjunto de sons A, B e C e supondo que o resultado que A foi melhor que B, e que B foi melhor que C, o ensaio será consistente se A for melhor que C e, inconsistente, se C for melhor que A. A repetitividade dos ensaios deverá ser maior ou igual a 70% e avaliadores com repetitividade menor que 60% devem ser desconsiderados. A média de consistência aceitável para cada jurado deve ser maior ou igual a 75%.

Uma vez que os pares tenham sido ajustados para atender aos requisitos acima descritos, o próximo passo consiste em analisar os dados. Uma maneira rápida de fazer a análise consiste em ordenar os sons de acordo com os escores obtidos nos julgamentos, sendo que o escore é dado pelo número de vezes que o som foi escolhido no julgamento. Essa análise é bastante rápida e simples, porém não fornece meios de se comparar com as métricas objetivas, uma vez que os escores dizem respeito ao como os sons estão em relação ao resto da população.

2.4.3 - *Ordenamento*

Os dados gerados pelo método de ordenamento indicam a posição relativa dos sons, mas não a magnitude das diferenças destes. Assim, os resultados deste método podem ser apresentados através do cálculo do posicionamento médio de cada som.

Para os testes de significância, o teste de Friedman é comumente utilizado [10]. O teste não paramétrico da menor diferença significativa de Fisher também pode ser utilizado.

Os testes de correlação mais utilizados são Kendal's Tau, Spearman Rank Correlation, e o Coeficiente de contingência sendo que destes, o último é o mais utilizado, tendo algumas vantagens em relação aos outros métodos não paramétricos, podendo-se citar que o Coeficiente de contingência não faz nenhuma suposição em relação ao tipo de distribuição, não exige nenhuma relação específica entre os dados e também pode ser estendido para outros tipos de situação [8].

CAPÍTULO 3 - REDES NEURAIS ARTIFICIAIS

Uma rede neural conforme HAYKIN [14] é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural de armazenar conhecimento experimental e torná-lo disponível para o uso. Ela se assemelha ao cérebro em dois aspectos:

1. o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem;
2. forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido;

O procedimento utilizado para realizar o processo de aprendizagem é chamado de algoritmo de aprendizagem, cuja função é modificar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo de projeto desejado.

A construção de uma RNA envolve os seguintes passos [15]:

1. determinação das propriedades da rede: a topologia (ou conectividade), o tipo de conexões, a ordem das conexões e a faixa dinâmica dos pesos;
2. determinação das propriedades dos neurônios que compõem a rede: a função de neurônio (ou função de transferência) e a faixa dinâmica de ativação dos neurônios;
3. determinação da dinâmica do sistema: envolve o esquema de inicialização dos parâmetros, o cálculo da função de cada neurônio e a regra de aprendizado.

3.1 - Topologia da rede

Um exemplo de modelo de RNA é apresentado na Figura 3.1. Este modelo é baseado em camadas, embora existam redes que não apresentem este tipo de estrutura.

Do ponto de vista de topologia, a forma mais comum de rede, é aquela organizada em camadas, as quais incluem:

- a camada de entrada;
- a camada oculta (uma ou mais de uma);
- a camada de saída;

Com relação à conectividade as redes podem ser:

- propagação para adiante (*feedforward*): todas as conexões apontam em uma única direção, da entrada para a saída;
- recorrentes: quando existem conexões de realimentação. Com relação ao tipo de conexão, podem ser do tipo simétrico ou assimétrico.

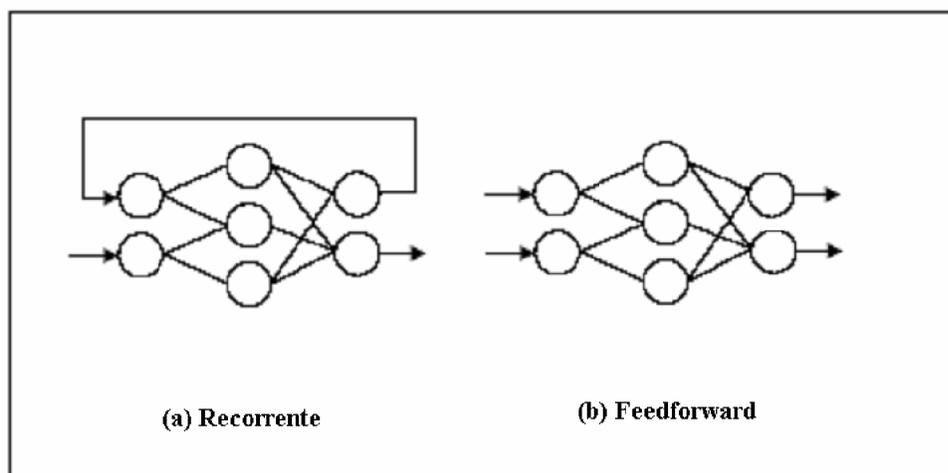


Figura 3.1 – Exemplo de modelo de uma RNA.

A ordem das conexões é determinada pelo número de saídas de neurônios, que são

combinadas em uma única conexão; tipicamente todas as RNAs são de primeira ordem.

Com relação aos pesos, estes podem ser números reais ou inteiros, e podem ou não ser confinados em uma determinada faixa. Após o treinamento os valores dos pesos não são, em geral, mais alterados.

3.1.1 - Funcionamento dos neurônios

A função de cada neurônio determina a ativação do mesmo. Esta função deve ser ajustável de forma a ficar compatível com a faixa dinâmica dos dados de entrada (dependente do problema), caso contrário poderá ocorrer do neurônio nunca ser ativado, ou permanecer sempre ativado.

3.1.2 - Dinâmica do sistema

A forma de inicialização dos parâmetros, depende do tipo de rede utilizado. No caso dos pesos das conexões, na maioria das vezes esta inicialização é feita com números pseudo-aleatórios pequenos. A regra de aprendizado é uma das características mais importantes de uma RNA. É a regra de aprendizado que determina como é feita a atualização dos parâmetros durante o treinamento, de forma a otimizar o desempenho da rede. Do ponto de vista de aprendizado, as RNAs podem ser apresentadas como de treinamento supervisionado ou não supervisionado. [16].

Entende-se por treinamento supervisionado, aquele no qual pares entrada/saída, representando a excitação e a resposta desejada (*target*), são sucessivamente apresentados à rede. O treinamento não supervisionado, por sua vez, consiste na apresentação apenas dos dados entrada para a rede, ficando a cargo desta a determinação das classes existentes.

3.2 - Características das RNAs

Uma das vantagens das RNAs é o processamento altamente paralelo que ocorre em sua estrutura. Existe um outro aspecto porém, o qual coloca as RNAs como uma classe muito especial de sistemas; é a capacidade de generalização exibida pelas RNAs. Por capacidade de generalização, entende-se a capacidade de gerar respostas adequadas, para entradas que nunca foram apresentadas durante o treinamento. Estes dois aspectos fazem com as RNAs sejam capazes de solucionar problemas altamente complexos.

Contudo, é importante observar que, na prática, as redes neurais artificiais não solucionam problemas sozinhas. Sua aplicação deve ser feita de forma criteriosa, decompondo o problema em questão em partes que possam ser tratadas eficientemente por uma ou várias RNAs, explorando ao máximo as habilidades de cada rede.

A utilização das RNAs, proporciona algumas características muito úteis aos sistemas que as utilizam [14,15], descritas a seguir.

3.2.1 - Não linearidade

A não linearidade é uma propriedade inerente ao neurônio, sendo extremamente importante nos casos onde o sinal (ou entrada) analisado é gerado por um dispositivo não linear, como é a maioria dos sistemas na natureza.

3.2.2 - Mapeamento entre a entrada e a saída

Um dos paradigmas de aprendizado mais utilizado é o aprendizado supervisionado, o qual promove a modificação dos pesos das conexões da rede, aplicando um conjunto de entradas com as respectivas saídas. Para cada uma destas entradas escolhida aleatoriamente, é

determinada a modificação dos pesos de modo a minimizar a diferença entre a resposta atual da rede, e a resposta desejada para a respectiva entrada. Desta forma, a rede aprende a partir de exemplos, construindo um mapeamento entre a entrada e a saída.

3.2.3 - Adaptatividade

As RNAs têm a capacidade natural de adaptar os seus parâmetros, em resposta às mudanças no ambiente a sua volta, através de um retreinamento. Contudo, se o ambiente é do tipo não estacionário, a RNA pode ser construída de forma a adaptar os parâmetros de suas conexões em tempo real, em resposta às mudanças no ambiente. Um cuidado porém deve ser observado quanto a esta adaptabilidade, uma vez que pode levar à perda de robustez da rede, ou seja, a RNA pode passar a responder a estímulos (ou entradas) aos quais não deveria responder, como por exemplo ruídos espúrios.

3.2.4 - Tratamento de informação contextual

Dependendo da arquitetura adotada, na fase de treinamento, a saída da rede é afetada pela atividade de todos os neurônios, e a correção dos parâmetros atinge potencialmente todas as conexões. Por exemplo, no caso dos pesos, este efeito faz com que a informação armazenada na rede seja influenciada por todo o conjunto de dados apresentado, considerando não apenas cada padrão apresentado mas também a correlação entre os mesmos, produzindo uma forma de informação contextual.

3.2.5 - Tolerância a falhas

Em vários modelos de rede, a saída recebe contribuições de todos os neurônios. Caso haja falha em um neurônio, ou em um conjunto de conexões, a saída será apenas degradada. As RNAs tendem a não apresentar falhas catastróficas com a perda de alguns neurônios, ou seja, são bastante tolerantes a falhas.

3.3 - Modelos de neurônios

A seguir é feita uma breve descrição de alguns modelos de neurônios clássicos mais conhecidos.

3.3.1 - O neurônio de McCulloch e Pitts

Os modelos de neurônios tiveram início com o trabalho de McCulloch e Pitts [17]. Em seu clássico artigo, descrevem um cálculo lógico de redes neurais que unificava os estudos de neurofisiologia e lógica matemática. Eles assumiam que o seu modelo formal de um neurônio seguia uma lei “tudo ou nada”. Com um número suficiente dessas unidades simples e com conexões sinápticas ajustadas apropriadamente e operando de forma síncrona, McCulloch e Pitts mostraram que uma rede assim constituída realizaria, a princípio, a computação de qualquer função computável. Este era um resultado muito significativo e com ele é geralmente aceito o nascimento das redes neurais e da inteligência artificial, [14].

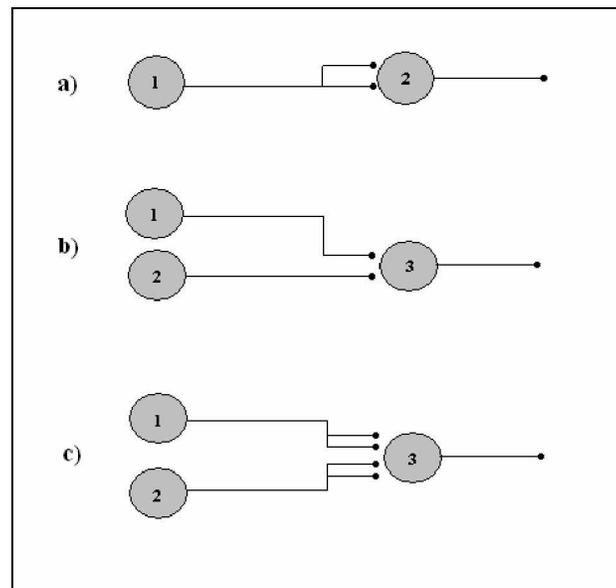


Figura 3.2 – Exemplo de redes simples de McCulloch-Pitts. O exemplo a) ilustra a lógica replicadora, b) a lógica E e c) a lógica OU. [18].

3.3.2 - O Perceptron

Um dos modelos de neurônio mais utilizados atualmente como bloco básico para a construção de redes mais complexas é o do perceptron [17]. A sua arquitetura é baseada no neurônio de McCulloch e Pitts, e a diferença com relação ao modelo anterior é o estabelecimento de um algoritmo de aprendizado com a utilização de pesos sinápticos [19], onde a saída é dada pela equação (3.1) e a atualização dos pesos é dada pela equação (3.2).

$$o(t) = F\left(\sum_0^{N-1} w_i \cdot i_i(t) - \theta\right), \quad (3.1)$$

$$w_i(t+1) = w_i(t) + \alpha \cdot e(t) \cdot i_i(t) \quad 0 \leq i \leq N-1, \quad (3.2)$$

onde: $e(t)$, corresponde ao erro, no caso de Rosenblatt $e(t) = d(t) - o(t)$;

α = ganho positivo < 1 , corresponde ao coeficiente de aprendizado.

Um exemplo bastante ilustrativo de como uma rede neural funciona como separador

de padrões é a classificação linear feita por uma rede de perceptrons, que é ilustrada no exemplo que a seguir.

Exemplo:

Supondo um perceptron com duas entradas (i_1 e i_2), deseja-se separar duas classes distintas A e B, descritas em função de i_1 e i_2 . A equação que descreve a saída fica:

$$o(t) = w_1 i_1(t) + w_2 i_2(t) + \theta. \quad (3.3)$$

A região limite de separação das classes, ocorre quando $o(t) = 0$, ou seja:

$$w_1 i_1(t) + w_2 i_2(t) + \theta = 0, \quad (3.4)$$

ou, ainda

$$i_2(t) = -\frac{w_1}{w_2} i_1(t) - \frac{\theta}{w_2}, \quad (3.5)$$

que é a equação de uma reta. A Figura 3.3, ilustra uma possível separação destes dois padrões. Deve ser observado, que até aqui, a função do neurônio corresponde ao degrau unitário (*hard limiting*), conforme descrito originalmente por Rosenblatt. Contudo na prática, a função sigmoidal do tipo:

$$o(t) = \frac{1}{1 + e^{-\beta \cdot x(t)}} \quad (3.6)$$

é mais utilizada, como não linearidade do perceptron, porque a ativação se suaviza. A função da Equação (3.6) é contínua. O fator β , corresponde à inclinação da sigmóide, e para valores de β muito grandes, a função tende para um degrau unitário (a função tangente hiperbólica também é muito utilizada).

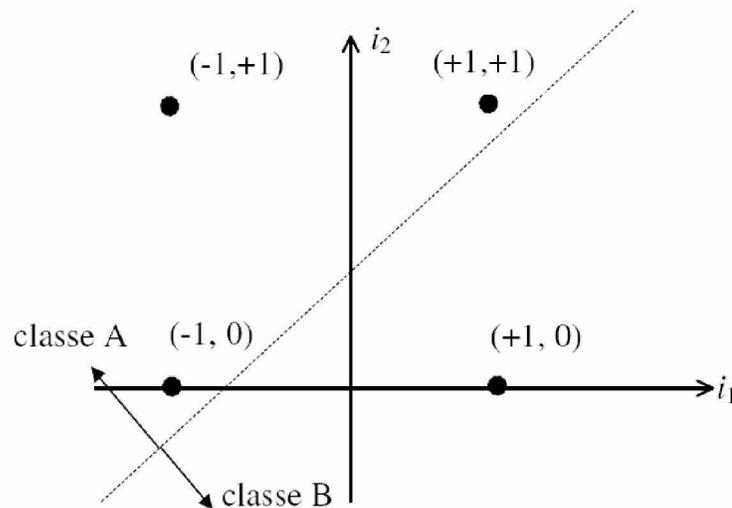


Figura 3.3 – Separação linear de dois padrões no espaço, por uma rede de dois perceptrons.

Algumas vantagens que justificam a utilização da função sigmóide, como função de ativação do neurônio, são apresentadas a seguir [19]:

- é uma função diferenciável, que permite a implementação de algoritmos de aprendizado do tipo gradiente descendente (*gradient descent*), para redes de múltiplas camadas, conforme será visto mais adiante neste trabalho;
- é adequada para aplicações que necessitem de uma saída com valores contínuos;
- produz uma saída com valores entre 0 e 1, que pode ser interpretada como uma probabilidade estimada, dependendo do tipo de rede e algoritmo de treinamento utilizado.

3.3.3 - O modelo do elemento linear adaptativo - Adaline

Outro modelo de neurônio importante é o Adaline, a diferença básica com relação ao

Perceptron está no fato da função binária de saída apresentar a seguinte forma:

$$F(x) = \begin{cases} +1 & \text{se } x \geq 0 \\ -1 & \text{se } x < 0 \end{cases} \quad (3.7)$$

e na adoção do erro médio quadrático (*LMS - Least Mean Square*), como procedimento para o cálculo do erro.

A equação para o cálculo do erro tem a forma dada a seguir:

$$e(t) = \left[\sum_i [o_i(t) - d_i(t)]^2 \right]^{1/2}, \quad (3.8)$$

o que para neurônios simples, conduz à mesma equação para adaptação dos pesos utilizada no perceptron.

A regra de aprendizado é conhecida como regra de Widrow, Hoff ou ainda regra delta e utiliza o método do gradiente descendente para a minimização do erro, sendo que os pesos são ajustados de forma proporcional ao negativo da derivada do erro com relação aos pesos anteriores [21,22].

3.3.4 - A RNA do tipo perceptron multicamadas ou *Multi Layer Perceptron*

A rede do tipo Perceptron Multicamadas ou *Multi Layer Perceptron (MLP)* é talvez a mais utilizada comercialmente. Esta é a rede utilizada nesta dissertação. Desta forma, é importante descrever o seu funcionamento. Conforme verificado anteriormente, o perceptron simples só é útil para a separação de classes linearmente separáveis, o que limita em muito a sua aplicação prática. A forma de contornar este problema é usar múltiplos perceptrons distribuídos em camadas totalmente conectadas. Na Figura 3.4 pode ser observado um exemplo de MLP com três camadas totalmente conectadas.

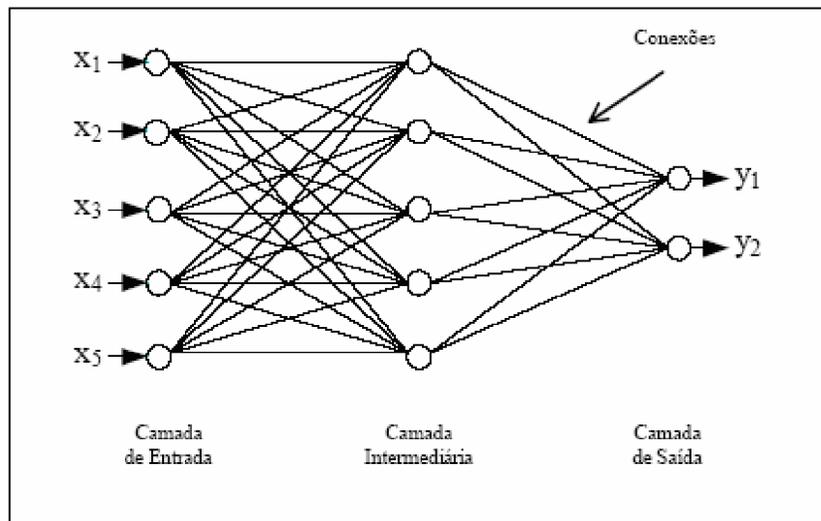


Figura 3.4 – Diagrama de um perceptron de múltiplas camadas de três camadas clássico.

Esta solução somente foi viabilizada com o desenvolvimento de um algoritmo de treinamento para redes com múltiplas camadas. O primeiro pesquisador a formalizar esta regra de aprendizado foi Werbos em 1974 [21], porém somente em 1986, com a divulgação do trabalho de Rumelhart [22], é que o algoritmo de retropropagação ou backpropagation tornou-se efetivamente conhecido.

3.3.5 - O algoritmo de retropropagação (backpropagation)

O algoritmo de backpropagation consiste na propagação do erro da saída da rede para a sua entrada. Este algoritmo foi desenvolvido a partir da generalização da regra delta, citada anteriormente neste trabalho, utilizando a técnica do gradiente descendente (*gradient descent*) para efetuar a correção dos pesos em função da derivada do erro com relação aos pesos anteriores [14]. As expressões genéricas para a atualização dos pesos, são as seguintes:

$$w(t+1) = w(t) + \Delta w(t) \quad (3.9)$$

para neurônios pertencentes a qualquer camada:

$$\Delta w(t) = \alpha \cdot \delta(t) i(t) \quad (3.10)$$

para um neurônio de saída:

$$\delta(t) = (d(t) - o(t)) \partial f(nf), \quad (3.11)$$

para um neurônio que não pertença a saída:

$$\delta(t) = \partial f(nf) \sum_k \delta_k(t) w_k(t), \quad (3.12)$$

onde: Δw é o valor da correção aplicada aos pesos

$\partial f(nf)$, é a derivada da função de ativação do neurônio: e

$\delta_k(t)$ e $w_k(t)$ são, respectivamente, a derivada da função de erro com relação aos pesos de um neurônio da camada subsequente e o peso da ligação entre o neurônio atual e o seu subsequente.

O termo $\delta(t)$ corresponde à derivada da função de erro com relação aos pesos, sendo responsável pela propagação do erro da saída da rede para as suas camadas intermediárias, através de um cálculo recursivo.

3.3.6 - O coeficiente de aprendizado α

Algumas considerações devem ser feitas sobre o coeficiente de aprendizado (α). Quando é aplicada a técnica do gradiente descendente, através do algoritmo de retropropagação, está sendo calculada uma trajetória no espaço de pesos em busca de um erro mínimo. Caso o valor adotado para α seja pequeno, a trajetória calculada será suave, porém o

tempo gasto para atingir-se o erro mínimo não será ótimo. Uma forma de compensar este fato é aumentar o valor de α de modo a acelerar o aprendizado da rede. Entretanto, este aumento no valor de α pode fazer com que a rede oscile e os valores dos pesos irão variar de forma abrupta, fazendo com que estes oscilem em torno do ponto de mínimo erro da rede. Uma maneira de contornar este problema é a adoção de um termo adicional para o cálculo da atualização dos pesos denominado momento (m) [14]. Este termo corresponde a um fator multiplicador, que determina um percentual da correção dos pesos da interação anterior que deve ser aplicado à correção atual, fazendo com que as oscilações sejam amortecidas pelo fator de inércia dos pesos assim criado. A aplicação deste recurso, altera a expressão para correção dos pesos para a seguinte forma:

$$\Delta w(t) = \alpha \delta(t) i(t) + m \Delta w(t-1) \quad (3.13)$$

A utilização do momento, porém, deve ser feita com critério, pois pode provocar o mesmo efeito de um coeficiente de aprendizado com valor muito alto. Um exemplo de outra solução, que é muito utilizada na prática, é o cálculo adaptativo do coeficiente de aprendizado e do termo correspondente ao momento.

3.3.7 - O problema do mínimo local

Um problema que ocorre com certa frequência, durante o treinamento da rede, é a convergência para um mínimo local da superfície de erro. Algumas formas para evitar este tipo de mínimo, são apresentadas em Hertz [19], as quais são citadas brevemente a seguir:

a) utilização do termo de momento, descrito anteriormente, fazendo com que sempre existam alterações mais adequadas nos pesos;

b) inicialização adequada dos pesos, evitando a inicialização com valores altos, o que faz com que o sistema fique “preso” em um mínimo local próximo do início do treinamento

(flat plateau);

c) uma solução alternativa para o item anterior é ajustar a função de transferência dos neurônios para a faixa dinâmica dos dados de treinamento, através de uma análise prévia;

d) adicionar ruído aos dados de treinamento, uma vez que estes mínimos em geral não são muito profundos, a correção nos pesos gerada por este ruído faz com que o sistema deixe o mínimo local. Isto pode ser conseguido através dos seguintes métodos:

- aplicação dos dados em seqüência aleatória;
- adicionar valores pequenos aleatoriamente aos pesos;
- somar valores pequenos aos dados de entrada. A quantidade de ruído adicionada, é um dado experimental, e seu uso incorreto pode causar instabilidade na rede.

3.3.8 - Otimização do Perceptron de múltiplas camadas

A seguir, são destacados os pontos considerados mais importantes da otimização do MLP:

a) A busca do tamanho ótimo da rede: este parâmetro tem influência direta na capacidade da rede construir um modelo adequado para um dado problema. Uma maneira de determinar o tamanho adequado da rede é considerar o número de exemplos de treinamento ou a dimensionalidade dos dados de entrada. Outra forma é utilizar técnicas para o crescimento ou diminuição seletiva do número de neurônios e conexões.

b) O tempo de treinamento da rede: uma vez determinado um tamanho adequado para a mesma, a minimização do tempo de treinamento pode ser feita através da escolha de coeficientes de aprendizado e técnicas para evitar a permanência da rede em mínimos locais durante o treinamento.

c) A capacidade de generalização da rede: consiste basicamente na habilidade da rede

produzir saídas significativas quando apresentada à padrões de entrada novos. A solução para este problema está intimamente ligada aos dois aspectos anteriores, sendo atingida através da utilização de um número suficiente de padrões de treinamento e da adequação da arquitetura da rede de forma a minimizar a representação de redundâncias e aspectos pouco significativos dos padrões representados.

CAPÍTULO 4 - METODOLOGIA DE AVALIAÇÃO SUBJETIVA

Como parte do desenvolvimento desta dissertação, chega-se à etapa fundamental para a avaliação dos conceitos e métricas da psicoacústica até então desenvolvidos, a realização de uma análise de júri. Esta é a base para qualquer trabalho que envolva a aplicação do conceito de qualidade sonora.

Com o objetivo de aplicar corretamente uma metodologia de avaliação subjetiva, anteriormente à realização dos experimentos foram tomados cuidados especiais, no que diz respeito à escolha e edição dos sinais a serem avaliados, ao local e à forma de apresentação dos ruídos, bem como com a escolha dos avaliadores que participaram do júri.

4.1 - Características gerais do compressor

Uma breve descrição do objeto de estudo deste trabalho, o compressor hermético de refrigeração, é apresentada neste item.

Os compressores herméticos são envoltos externamente por uma carcaça, formada por duas partes separadas, corpo e tampa, que permitem a instalação do conjunto interno. Ao final da montagem a tampa é soldada ao corpo. Fixados ao corpo encontram-se a placa base, o terminal hermético e os passadores de sucção, de descarga e de processo. A função da placa base é fixar o compressor ao sistema de refrigeração, enquanto o terminal hermético fixado à carcaça faz a conexão elétrica entre o conjunto e a rede elétrica.

Os componentes internos dividem-se em dois grupos principais: parte mecânica, o compressor, e a parte elétrica, o motor de indução, montados sobrepostos e ligados por um eixo. O conjunto moto-compressor é apoiado sobre molas. O óleo lubrificante, depositado no

fundo da carcaça, é levado às partes superiores do compressor por meio de uma bomba fixada ao rotor.

O gás de refrigeração ocupa o espaço formado entre a carcaça e o conjunto motor-compressor, que é succionado para o interior do cilindro através do filtro de sucção. Este possui uma importante função, a de atenuador acústico.

4.2 - Sinais de ruído

Dentre os diversos produtos disponíveis na empresa, foram concedidos para a realização deste trabalho os dados de ruído de um compressor em fase de desenvolvimento. Este, foi medido e monitorado durante um período de aproximadamente quatro meses.

Os sinais de ruído utilizados foram gravados em um ambiente de campo difuso que representa a resposta de cozinhas típicas [2], a uma distância de um metro e cinquenta centímetros da fonte sonora, por um microfone fixo, ou seja, captação monoaural. Por esta razão fez-se necessária a avaliação no que diz respeito à qualidade e a fidelidade dos sinais gravados, já que é recomendável a utilização de um conjunto simulador cabeça e tronco. A função deste é simular a presença do corpo humano no campo sonoro. Este simulador possui um par de microfones colocados na posição dos ouvidos, dita captação binaural. Os fatores que determinam a diferença entre as duas formas de captação do sinal para a gravação são, principalmente, a consideração das distorções causadas pela presença do corpo humano no campo acústico e a possibilidade da determinação da posição da fonte sonora. Entretanto, na impossibilidade de utilizar o conjunto simulador cabeça e tronco, utilizarem-se os sinais gravados com microfone aplicando-se filtros de correções disponíveis no software de qualidade sonora, para que seja levada em consideração a interferência da presença de pessoas no local de gravação.

4.3 - Local de teste

Levando-se em conta a dificuldade de se realizar as análises de júri com a apresentação dos ruídos diretamente, optou-se pela reprodução digital. A reprodução de sinais digitais pode ser feita por caixas de som em ambientes reverberantes ou fones de ouvido. A vantagem do uso de fones de ouvido em experimentos psicoacústicos é de apresentarem uma pequena distorção não linear, menor que 0,1 % na faixa de frequência de interesse. Apresentam, também, vantagens quanto ao local de reprodução e quanto ao ruído de fundo, sofrendo uma menor influência. Escolheu-se assim a reprodução com fones de ouvido.

Recapitulando do Capítulo 2, deve-se lembrar do primeiro requisito para o ambiente de teste é o baixo valor do nível de ruído de fundo. Assim foi alocada uma sala de testes, localizada no laboratório de vibrações e acústica da Embraco em Joinville, que apresentasse ou se aproximasse de tal característica, isso para que o mesmo não interfira no julgamento dos jurados. Foram tomados cuidados especiais com possíveis fontes de ruído, tais como computadores, lâmpadas fluorescentes e outros, para que não afetem os resultados dos testes.

A sala utilizada foi construída e qualificada para medições desta natureza, facilitando assim a realização do experimento, já que um número reduzido de pessoas tem acesso à mesma.

4.4 - Seleção dos jurados

Considerando a importância da familiaridade com o produto, a fim de evitar que suas expectativas em relação ao produto influenciem o julgamento, os jurados devem ser escolhidos de tal forma que as mesmas não comprometam os resultados dos ensaios. Assim,

como é comum em empresas que realizam testes de júri, os jurados foram voluntários da Embraco, uma vez que estes possuem um conhecimento sobre os produtos testados.

A experiência dos jurados, com os ensaios, também pode ser um fator importante. Além do fato de jurados mais experientes conseguirem perceber com maior clareza algumas nuances dos sons apresentados, também deve ser considerado, por outro lado, que peritos no assunto tendem a valorizar aspectos que podem não ser importantes ao usuário final. Assim, procurou-se efetuar os testes com pessoas com alguma familiaridade com os ruídos e com pouca experiência em ensaios de júri.

Por fim, buscou-se realizar o maior número de avaliações possíveis. Contudo conseguiu-se o número de 26 jurados, entre funcionários e pesquisadores envolvidos com a empresa. Este número deve ser levado em conta, pois está diretamente relacionado à representatividade dos resultados.

4.5 – Ajuste de ganho dos sinais

Para a apresentação de sinais digitais em uma análise de júri, seja por caixas acústicas ou fones de ouvido, é de grande importância que os ruídos reproduzidos estejam com níveis de amplitude semelhantes aos de origem. No caso da reprodução destes sinais, seja com ampliações ou atenuações, a percepção dos avaliadores é diferente, resultando em avaliações que não representam a situação real de funcionamento dos compressores. Para corrigir tal situação, os sinais de reprodução foram ajustados da seguinte forma:

1. utilizando-se do calibrador modelo, Brüel & Kjaer tipo 4294, gravou-se o sinal de referência, com um nível de pressão sonora de 74 dB, no ambiente em que foram adquiridos os sinais dos compressores;

2. posicionou-se um microfone binaural, modelo Brüel & Kjaer tipo 4101, no sistema auditivo de um jurado, e reproduziu-se o sinal de referência através do fone de ouvido, modelo SENNHEISER[®] HD 580 precision, ajustando-se o ganho para que o mesmo reproduzisse os 74 dB medidos. Na Figura 4.1 é apresentado o modelo de microfone binaural utilizado no processo de ajuste.



Figura 4.1 – Microfone binaural utilizado no processo de ajuste.

4.6 - Apresentação dos sinais

A forma de apresentação dos sinais tem como objetivo implementar a qualidade e confiabilidade das informações geradas. Definidos o ambiente de teste e meio de apresentação dos sinais, via fones de ouvido de uso profissional (SENNHEISER[®] HD 580 precision), o ponto passa a ser a forma de apresentação ao júri. Dentre os métodos citados na revisão bibliográfica apresentada no capítulo dois, o método de comparação pareada se mostrou mais atraente, levando-se em conta as características e as informações desejadas neste ensaio e o número de ruídos avaliados.

Uma pequena alteração no método foi realizada, a fim de que o tempo de ensaio fosse reduzido significativamente. Para que os jurados não precisassem avaliar pareadamente todos os ruídos disponíveis, uma referência foi escolhida. Assim os sinais restantes foram comparados com esta referência.

A avaliação dos sinais também foi restringida à escolha das seguintes classes:

- ruídos mais agradáveis do que a referência;
- ruídos semelhantes à referência;
- ruídos mais incômodos do que a referência;

A explicação desta restrição, três classes de escolha, se encontra na necessidade de se obter um grande número de amostras por classe, para o treinamento adequado de uma rede neural artificial. Assim é fácil deduzir que ao aumentar o número de classes de escolha, o número de amostras por classe diminui. Por exemplo, aumentando-se o número de classes de três para cinco, o número de exemplos por classe passaria de aproximadamente trinta por classe para cerca de dezesseis por classe. Para tanto seria necessário que o número de amostras passasse da ordem dos noventa para a ordem dos cento e cinquenta ruídos.

4.7 - A análise de júri

Este experimento baseia-se, primeiramente, na entrega de um formulário para o avaliador, possibilitando-o a fazer as marcações que lhe forem solicitadas. São passadas ao jurado as informações necessárias para a realização do teste. Um conjunto de ruídos exemplos é amostrado, de forma que o avaliador possa formar um escala subjetiva dos sons a serem

avaliados. Não havendo necessidade de maiores esclarecimentos por parte do jurado, o experimento é iniciado, sendo que a apresentação dos sinais feita da seguinte maneira:

1. Ruído de referência
2. Ruído do compressor 1
- *Pausa de um segundo*
3. Ruído de referência
4. Ruído do compressor 2
- *Pausa de um segundo*
5. Ruído de referência
6. Ruído do compressor 3
- *Pausa de um segundo*

:

:

183. Ruído de referência
184. Ruído do compressor 92

Finalizando assim com o ruído do compressor de número noventa e dois.

Quanto à duração do teste, deve-se considerar a duração dos sinais (dez segundos), a quantidade de sinais (noventa e dois), e o fato de que, antecedendo o próximo som a ser avaliado, é rerepresentado o som de referência. Em resumo, o teste tem uma duração aproximada de trinta e dois minutos. A duração do experimento, apesar de estar dentro dos limites de tempo recomendados, por volta de 40 minutos, foi considerada cansativa por aqueles que participaram do teste. Este é um fator importante, que também deve ser levado em consideração, pois as características dos ruídos influem diretamente na receptividade dos jurados. Segundo NUNES [2], os jurados aceitam, por exemplo, com muito mais facilidade

um número maior de ruído de motores de automóveis do que ruídos de diferentes secadores de cabelos, mesmo tendo avaliado em ambos os grupos de ruídos extremamente desagradáveis. Um exemplo da tabela utilizada é mostrado na Figura 4.4.

	Mais agradável	Semelhante	Mais incômodo
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
89	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
90	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
91	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
92	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Figura 4.2 – Modelo de tabela apresentada aos jurados para seleção dos ruídos apresentados.

4.8 - Resultados da análise de júri

Com o objetivo de gerar um banco de dados que tenha um caráter generalista, não foi descartado nenhum resultado de análise subjetiva, pois todos apresentaram-se coerentes com suas avaliações. Tal fato foi verificado com a repetição de alguns dos sinais, bem como da própria referência.

Do total de 92 ruídos, a distribuição resultante das classes está resumida na Tabela 4.1.

Tabela 4.1 – Distribuição das classes de ruído conforme análise de júri.

Classe	Nº de sons
Ruídos considerados mais agradáveis do que a referência	31
Ruídos considerados semelhantes à referência	30
Ruídos considerados mais incômodos do que a referência	31

É interessante observar a distribuição dos ruídos, de forma que, mesmo tendo-se feito a escolha do ruído referência, a distribuição ficou uniforme, ao considerar-se que se tratam de ruídos de vários lotes de compressores.

Nas Figuras 4.5, 4.6 e 4.7 são apresentados os resultados das avaliações, estando os ruídos divididos em suas respectivas classes. Os gráficos indicam o percentual de escolha para o som analisado. Lembrando que o percentual de 100% significa a escolha de 26 jurados.

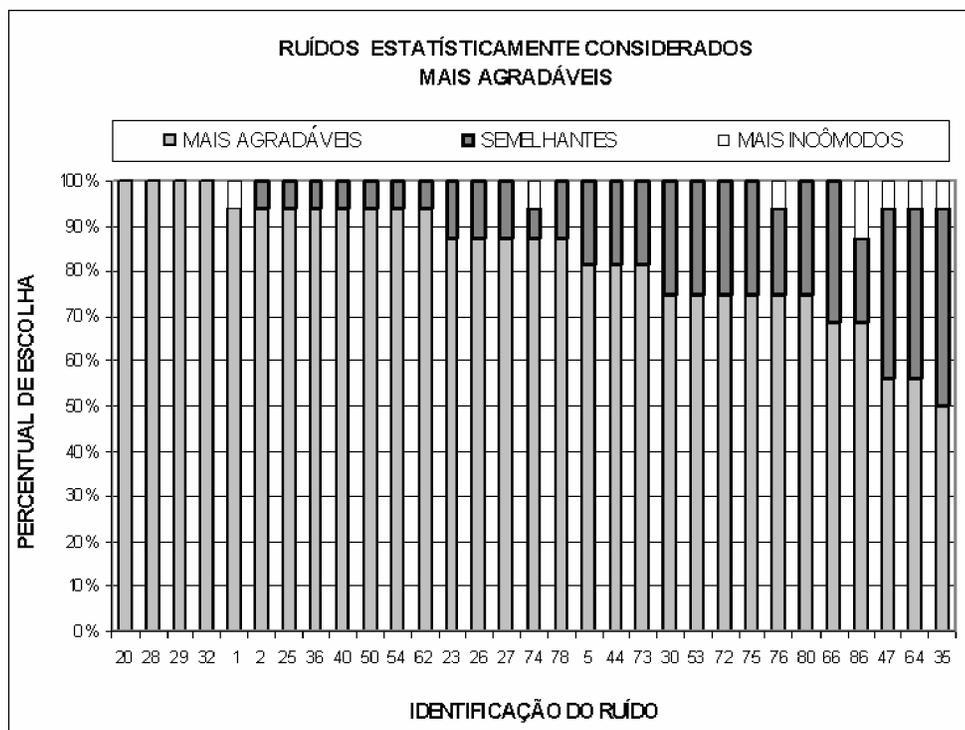


Figura 4.3 – Distribuição das avaliações subjetivas para os sinais considerados mais agradáveis.

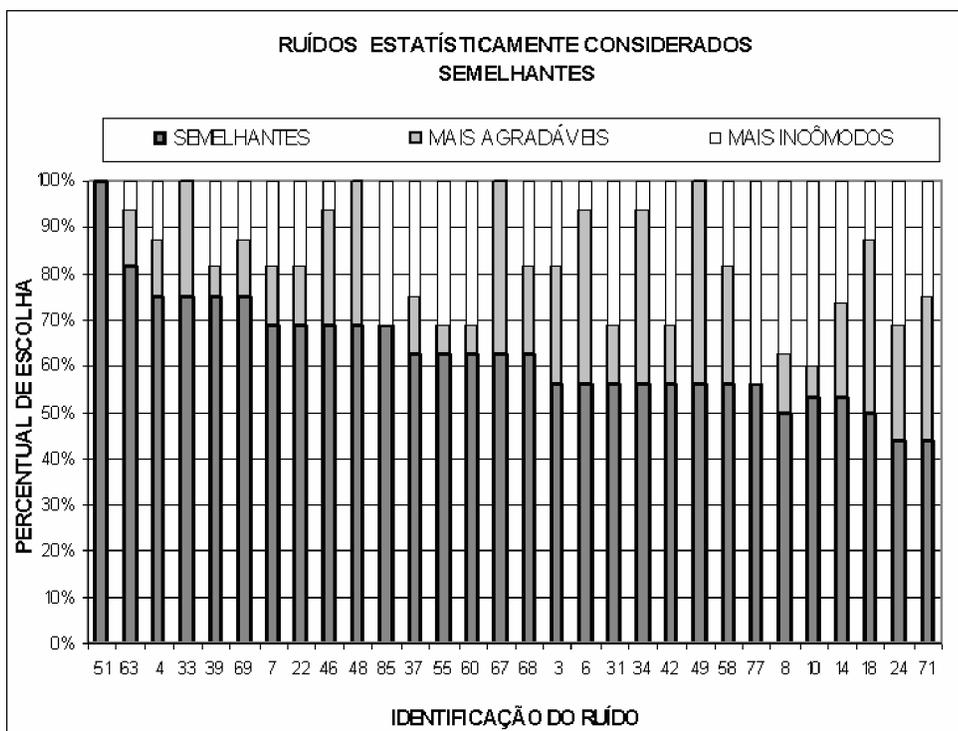


Figura 4.4 – Distribuição das avaliações subjetivas para os sinais considerados semelhantes.

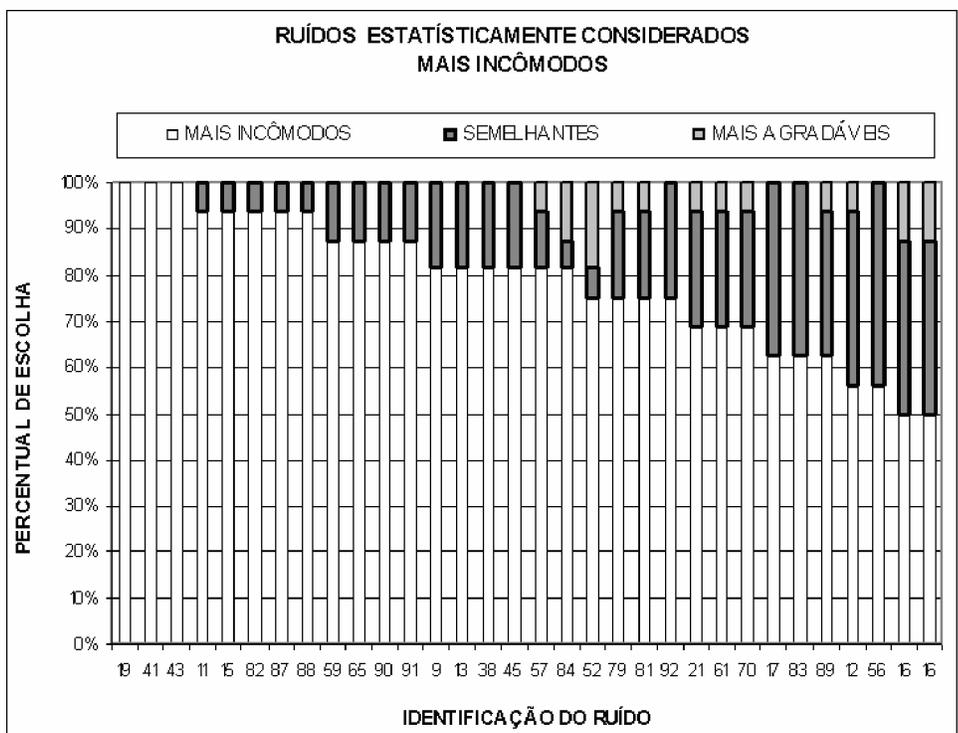


Figura 4.5 – Distribuição das avaliações subjetivas para os sinais considerados mais incômodos.

Antes de fazer uma avaliação dos resultados apresentados pela análise de júri é interessante que se comentem alguns pontos importantes, identificados nesta etapa do trabalho e que alteram as avaliações subjetivas dos jurados. Refere-se ao ganho aplicado aos sinais e à sua ordem de apresentação. Constatou-se que os jurados tendem a avaliar de forma diferente os mesmos sinais, quando estes são apresentados em ordem diferente, principalmente no caso dos ruídos com sonoridades muito diferentes. Também foi verificado uma diferença de percepção para diferentes ganhos dos sinais.

A análise dos resultados mostra que são poucos os ruído classificados por unanimidade do júri, sendo que o percentual de escolha dos demais sons mostram claramente a dispersão de opiniões. Deve-se ter em mente que esta é uma análise subjetiva, portanto resultados como este são esperados.

Observa-se também a menor divergência de opiniões para a classificação dos ruídos mais agradáveis e dos menos agradáveis, estando as maiores indecisões nos ruídos considerados semelhantes. Isto pode ser explicado pelo fato de que, quando os níveis dos sinais são semelhantes, características particulares são levadas em consideração, o que não ocorre nos extremos, muito pior ou muito melhor.

Por fim vale ressaltar algumas vantagens da apresentação digital, aqui utilizada. A possibilidade de se reproduzirem os testes em outros locais, que não o local de gravação, a garantia de que os jurados estarão avaliando os mesmos ruídos e a menor interferência do ruído de fundo.

Deve-se ainda salientar que a incerteza destes resultados terá impacto direto na capacidade classificação da rede neural artificial, evidenciando a importância de uma análise de júri criteriosa.

CAPÍTULO 5 - BASE DE DADOS PARA A REDE NEURAL

A etapa que antecede a utilização de uma rede neural artificial é a elaboração de um banco de dados, sendo este um fator determinante para a utilização bem sucedida desta ferramenta. É apresentada, a seguir, a metodologia utilizada para a configuração desta base, descrevendo as etapas envolvidas no processo.

Uma descrição geral dos dados coletados é apresentada.

5.1 - A elaboração do banco de dados

A base de dados utilizada neste trabalho foi montada a partir da correlação das avaliações realizadas por uma análise de júri com valores discretos destes ruídos. Uma vez adquiridos os noventa e dois sons na forma de pressão sonora no domínio do tempo, foi realizado o processamento destes sinais, passando-os para o domínio da frequência, obtendo assim espectros de ruído em terço de oitava. Processamentos adicionais foram realizados até os valores finais destes, chegando-se aos seguintes valores discretizados: espectro dB (A), curva de *loudness*, curva de *fluctuation strength*, curva de *roughness* e valores globais de *sharpness* e *prominence ratio*.

A Figura 5.1 ilustra o fluxograma dos processos envolvidos na configuração da base de dados.

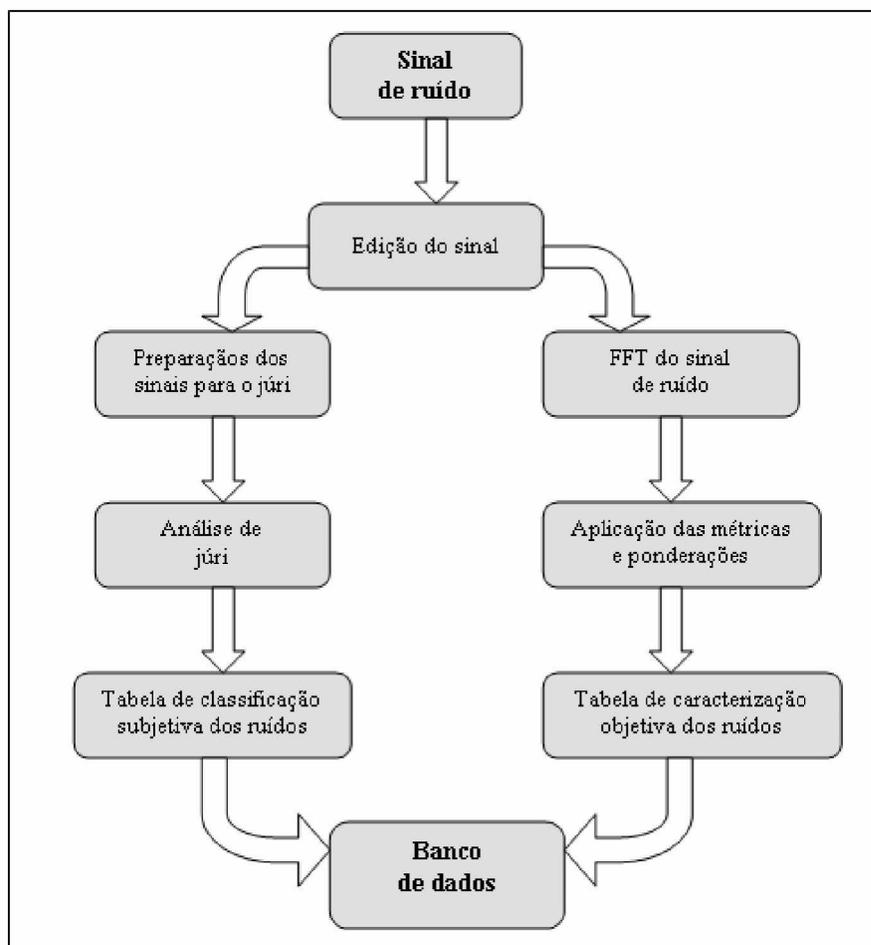


Figura 5.1 – Fluxograma de elaboração do banco de dados.

5.2 - Dados de classificação subjetiva do ruído

O procedimento utilizado para a avaliação subjetiva está descrito no Capítulo 4, sendo portando desnecessária sua rerepresentação. O item relevante para esta etapa é o conjunto de resultados obtidos naquele experimento.

A informação relevante neste momento se resume à classificação subjetiva dos ruídos avaliados, levando-se em conta a maioria das opiniões assinaladas. Isto, pelo fato de em um primeiro instante não ser avaliado o grau de certeza que um determinado som pertence àquela

população de sons. A Tabela 5.1 apresenta a distribuição de votos dos ruídos considerados mais agradáveis.

Tabela 5.1 – Tabela de distribuição de votos para os sons considerados mais agradáveis

	Mais agradáveis nº de votos	Semelhantes nº de votos	Mais incômodos nº de votos	Mais agradáveis %de votos	Semelhantes % de votos	Mais incômodos %de votos
20	25	0	0	100.0	0.0	0.0
28	25	0	0	100.0	0.0	0.0
29	25	0	0	100.0	0.0	0.0
32	25	0	0	100.0	0.0	0.0
1	23	1	1	92.0	4.0	4.0
2	23	2	0	92.0	8.0	0.0
25	23	2	0	92.0	8.0	0.0
36	23	2	0	92.0	8.0	0.0
40	23	2	0	92.0	8.0	0.0
50	22	3	0	88.0	12.0	0.0
54	22	3	0	88.0	12.0	0.0
62	22	3	0	88.0	12.0	0.0
23	21	3	1	84.0	12.0	4.0
26	21	2	2	84.0	8.0	8.0
27	21	2	2	84.0	8.0	8.0
74	21	3	1	84.0	12.0	4.0
78	21	4	0	84.0	16.0	0.0
5	20	5	0	80.0	20.0	0.0
44	20	5	0	80.0	20.0	0.0
73	20	5	0	80.0	20.0	0.0
30	18	7	0	72.0	28.0	0.0
53	18	7	0	72.0	28.0	0.0
72	18	7	0	72.0	28.0	0.0
75	18	7	0	72.0	28.0	0.0
76	18	5	2	72.0	20.0	8.0
80	18	7	0	72.0	28.0	0.0
66	17	8	0	68.0	32.0	0.0
86	17	5	3	68.0	20.0	12.0
47	14	10	1	56.0	40.0	4.0
64	14	10	1	56.0	40.0	4.0
35	13	10	2	52.0	40.0	8.0

Um dos exemplos da questão abordada pode ser o ruído número setenta e seis, que apesar de ser considerado mais agradável pela maioria dos jurados, foi votado como ruído semelhante, e também mais incômodo, tendo respectivamente, setenta e dois por cento, vinte por cento e oito por cento do total de votos.

A formatação da classificação dos ruídos com critérios subjetivos conclui esta primeira etapa da configuração do banco de dados. Assim, cada som avaliado foi classificado por maioria de votos, não havendo casos de indecisão.

5.3 - Dados de caracterização objetiva do ruído

Os sinais de ruído utilizados foram adquiridos pelo sistema de medição Brüel & Kjaer, compostos por um módulo analisador de sinais de 32 canais, um microfone de uma polegada, tipo 4145, software de aquisição de sinais PULSE v.8.0 também do fabricante B&K e cabos de conexão, disponibilizados pela Embraco, em Joinville, SC. Esta etapa do processamento destes sinais foi inteiramente realizada no software de qualidade sonora, também da Brüel & Kjaer. Um sinal típico de pressão sonora no domínio do tempo utilizada para os processamentos é apresentado na Figura 5.2.

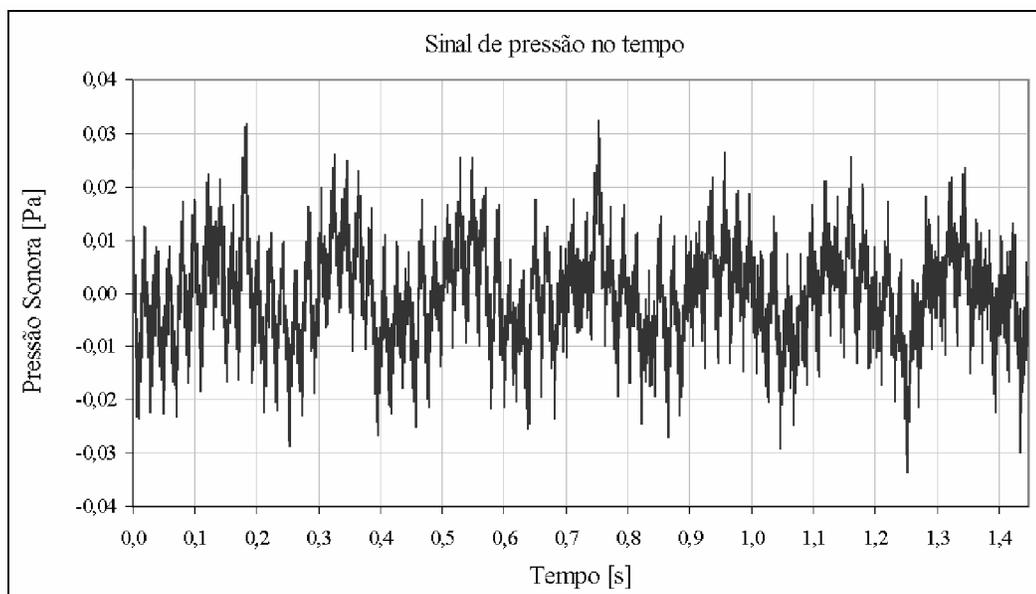


Figura 5.2 – Sinal de ruído no domínio do tempo.

O trabalho realizado por Nunes [23], no qual avaliou a aplicação de algumas métricas de qualidade sonora disponíveis no aplicativo Sound Quality do fabricante Brüel & Kjaer, em

ruídos de compressores herméticos e sistemas de refrigeração, concluiu que os valores que obtiveram boa correlação com o júri realizado em seu trabalho, foram a métrica de *loudness* e a curva de ponderação A. Desta forma, neste trabalho foram também considerados em um primeiro momento, os valores em dB (A) e os valores da métrica de *loudness* dos sinais de ruído.

A curva de ponderação A é a maneira mais comum de considerar aspectos subjetivos, a qual considera a característica de não linearidade da sensibilidade do sistema auditivo. O *loudness*, por ser uma métrica que representa de forma mais precisa a sonoridade do ruído também foi utilizado como dado de entrada.

As outras métricas disponíveis, como o *sharpness*, que descreve o equilíbrio espectral relacionando à distribuição de energia entre altas e baixas frequências, o *roughness*, que caracteriza modulações de médias frequências do sinal, o *fluctuation strength* que caracteriza modulações lentas e por fim o *prominence ratio* como forma de avaliar os tons puros presentes nos sinais, foram avaliadas em um segundo momento deste trabalho. A Figura 5.2 mostra um espectro de valores em dB (A) típicos dos sinais utilizados.

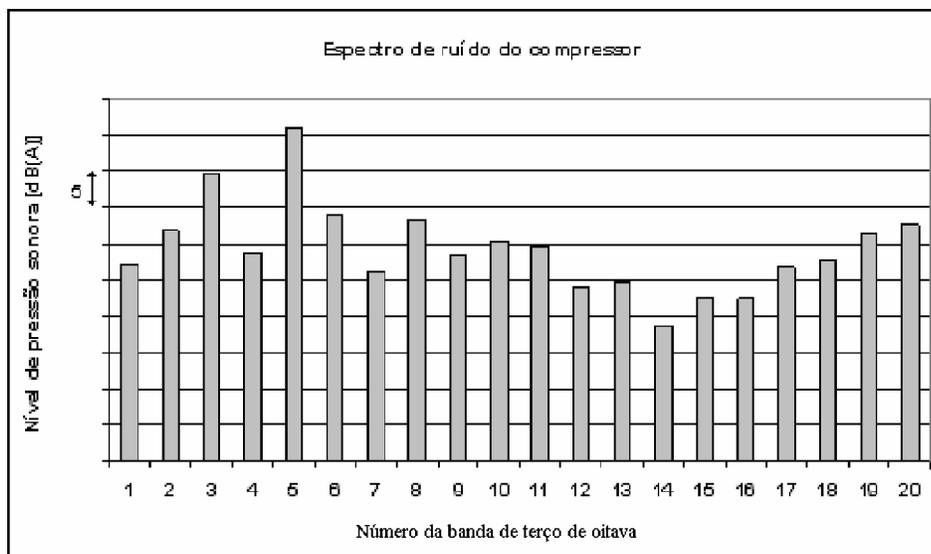


Figura 5.3 – Espectro de um terço de oitava utilizado como dado de entrada para a rede neural.

Valores típicos de *loudness*, *fluctuation strenght* e *roughness* também são apresentados nas Figuras 5.3, 5.4 e 5.5. Notam-se os valores em escala de banda crítica, conceito apresentado no Capítulo 2.

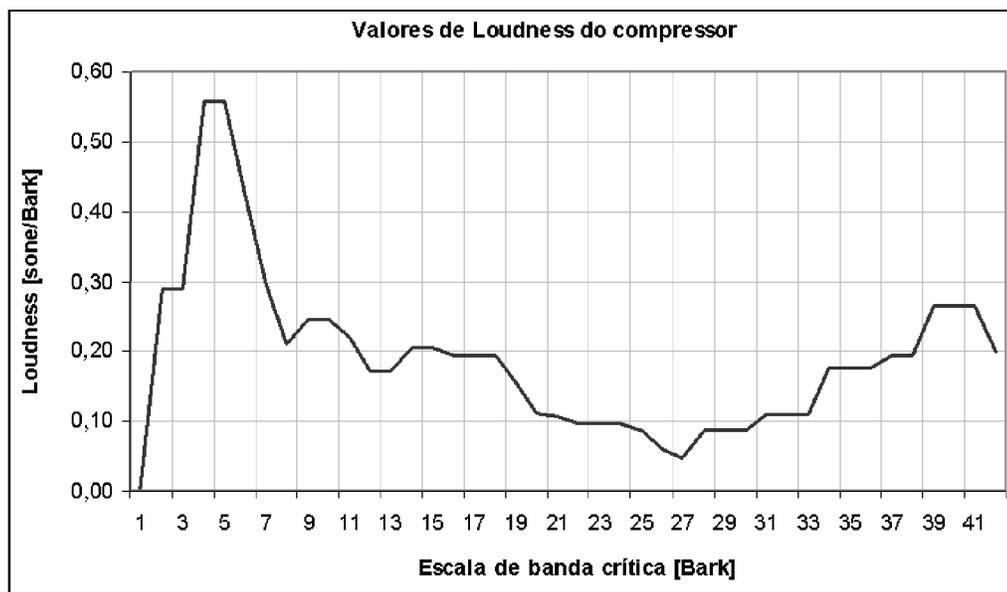


Figura 5.4 – Valores de Loudness utilizados como dado de entrada para a rede neural.

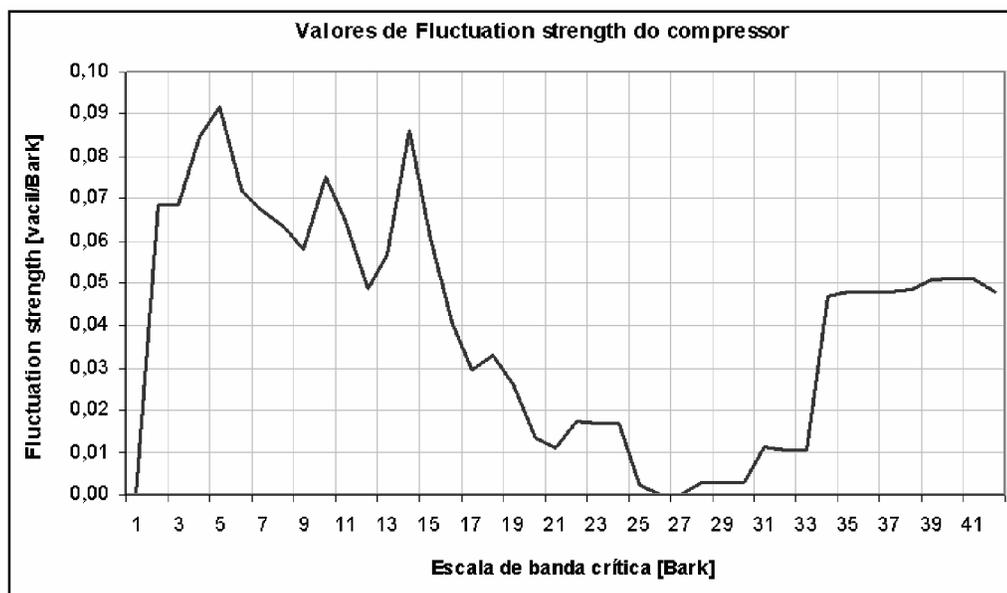


Figura 5.5 – Valores de fluctuation strength utilizados como dado de entrada para a rede neural.

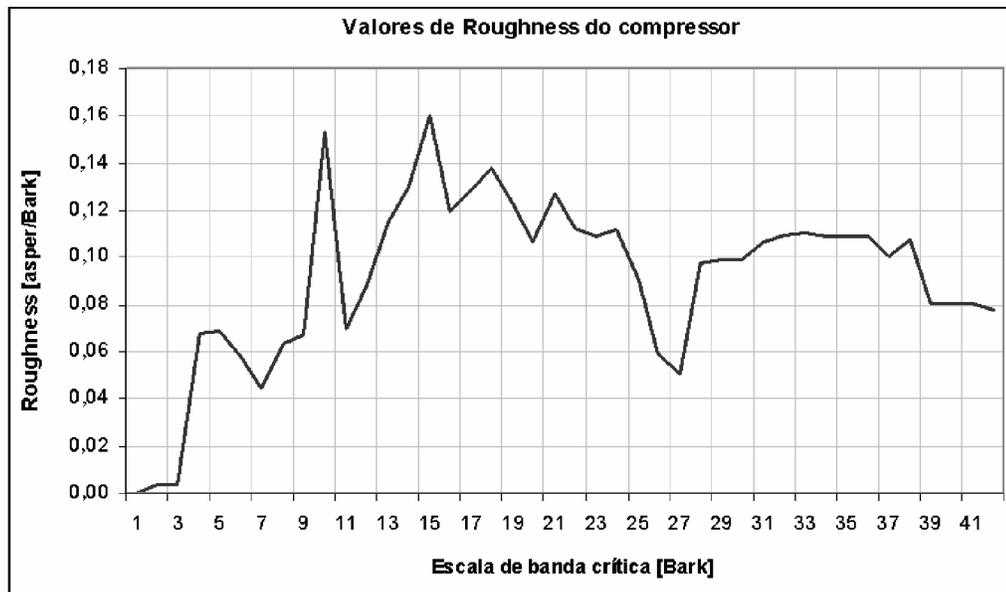


Figura 5.6 – Valores de roughness utilizados como dado de entrada para a rede neural.

A apresentação destes valores, que a partir de então passarão a se chamar atributos, em uma formatação adequada para a utilização da rede neural, é descrita na seqüência.

5.4 - A base de dados

A formação da base de dados, conforme indica o fluxograma da Figura 5.1, consiste na união das informações subjetivas, obtidas através da análise de júri, com as informações ditas objetivas, os valores processados, originados de ponderações e de conceitos psicoacústicos. Uma ilustração da base de dados utilizada é apresentada na Figura 5.6. Os valores estão normalizados.

	Som 1	Som 2	Som 3	Som 4	Som 91	Som 92
Stationary Loudness	0.537	0.686	0.645	0.777	0.686	0.504
Loudness Médio	1.000	0.678	0.733	0.744	0.811	0.700
Inst.Loudness Médio	0.613	0.442	0.270	0.749	0.800	0.638
dBA global	0.284	0.180	0.415	0.583	0.398	0.322
dBA 1	0.408	0.354	0.354	0.669	0.677	0.531
dBA 2	0.627	0.377	0.435	0.366	0.321	0.506
dBA 3	0.343	0.182	0.090	0.329	0.659	0.635
dBA 4	0.457	0.308	0.170	0.447	0.887	0.861
dBA 5	0.537	0.386	0.343	0.288	0.482	0.515
...
...
...
...
dBA 15	0.418	0.280	0.454	0.803	0.626	0.526
dBA 16	0.744	0.409	0.550	0.934	0.761	0.662
dBA 17	0.831	0.706	0.753	0.747	0.930	0.842
dBA 18	0.849	0.644	0.641	0.736	0.928	0.905
dBA 19	0.682	0.517	0.422	0.742	0.742	0.705
dBA 20	0.744	0.409	0.550	0.934	0.761	0.662
Loudness 1	0.392	0.657	0.332	0.436	0.517	0.427
Loudness 2	0.360	0.385	0.411	0.503	0.536	0.542
Loudness 3	0.325	0.366	0.261	0.726	0.623	0.490
Loudness 4	0.260	0.383	0.287	0.663	0.710	0.432
Loudness 5	0.090	0.276	0.260	0.471	0.661	0.405
...
...
...
...
Loudness 38	0.091	0.091	0.071	0.099	0.237	0.222
Loudness 39	0.106	0.034	0.082	0.230	0.292	0.273
Loudness 40	0.103	0.104	0.087	0.064	0.207	0.201
Loudness 41	0.179	0.176	0.033	0.505	0.247	0.205
Loudness 42	0.042	0.236	0.058	0.366	0.200	0.219
Classe	M. Agradável	M. Agradável	Semelhante	M. Incômodo	Semelhante	M. Incômodo

Figura 5.7 – Formato da tabela utilizada como banco de dados para a rede neural artificial.

Com um total de sessenta e seis atributos e três classes de saída está formado o banco de dados para ser utilizado na configuração da RNA. A matriz de dados tem um tamanho final de sessenta e sete linhas e noventa e duas colunas, sendo necessário agora, apenas a transposição dessa matriz e a normalização dos dados para a forma final do conjunto de treinamento da RNA.

5.5 - Descrição geral da base de dados

Uma descrição resumida da base de dados é apresentada na Tabela 5.2.

A questão do erro majoritário, até então não abordada neste trabalho, refere-se ao menor erro obtido ao tentar-se selecionar aleatoriamente um indivíduo pertencente à maior

classe de classificação. Outro esclarecimento se refere aos atributos, ou as entradas, sendo todas números reais, isto é, contínuos, havendo apenas dados nominais nas classes, ou seja, nas saídas dos neurônios.

Tabela 5.2 – Tabela descritiva da base de dados

Conjunto de dados	Exemplos	Duplicados ou conflitantes	Atributos (contínuos / nominais)	Classes	Tamanho das classes	Erro majoritário	Valores desconhecidos
Jury	92	0 (0 %)	154 (154 / 0)	Mais agradáveis Semelhantes Mais incômodos	33,70% 32,60% 33,70%	66,30%	Não

O erro majoritário é um parâmetro indicador da efetividade da rede neural, isto é, o percentual de erro da rede neural deve estar abaixo deste; desta forma a rede estará contribuindo para a classificação dos dados. O caso da situação inversa, erro da rede maior do que o erro majoritário, indica que a RNA não adquiriu nenhum aprendizado, sendo descartada a sua utilização como classificador.

CAPÍTULO 6 - CONFIGURAÇÃO DA REDE NEURAL ARTIFICIAL

O modelo de neurônio a ser utilizado neste trabalho, conforme citado na revisão sobre redes neurais apresentada no Capítulo 2, é o *Perceptron* de múltiplas camadas, ou o *Multi Layer Perceptron*. A seqüência deste capítulo envolve a apresentação da ferramenta utilizada para a configuração de uma RNA de classificação e a metodologia utilizada neste desenvolvimento. Também é feita uma análise do banco de dados utilizado, verificando-se a relevância dos atributos da rede, ou seja, dos valores de entrada que geraram aprendizado para a rede. Esta análise final permite concluir quais as métricas e faixas de frequência que melhor representam a classificação da análise de júri, no que diz respeito aos ruídos do modelo de compressor utilizado nesta dissertação.

6.1 - Ferramenta utilizada

Para avaliação, elaboração e configuração da rede neural artificial neste trabalho, utilizou-se como ferramenta computacional o programa Weka (*Waikato Environment for Knowledge Analysis*), este, um software de comercialização livre desenvolvido pela Universidade de Waikato, Nova Zelândia. Implementado em Java, tem como principal característica ser portátil, isto é, pode rodar nas mais variadas plataformas, aproveitando-se, ainda, dos benefícios de uma linguagem orientada a objetos. Além disso, o software contém ferramentas de pré-processamento de dados, classificação, regressão, associação de regras e visualização e pode ser aplicado diretamente a uma base de dados ou por código em Java.

Dentre os métodos de aprendizado de máquina disponíveis o software possui o implementado o algoritmo de rede neural artificial *multilayer perceptron* com aprendizado por *backpropagation* que é utilizado neste trabalho.

6.1.1 - Formatação do arquivo de relações e atributos – ARFF

Para se utilizar o software Weka é necessário que a base de dados esteja no formato ARFF (*Attribute-Relation File Format*), o qual descreve o domínio dos atributos e a classe do mesmo.

A formatação consiste apenas na elaboração de um cabeçalho que indica esses parâmetros (definindo se seus valores são contínuos ou nominais) e, posteriormente, a este adicionam-se os dados. Um exemplo de base de dados no formato ARFF pode ser observado na Figura 6.1.

```

@RELATION jury
@ATTRIBUTE Stationary_Loudnes REAL
@ATTRIBUTE Loudness_Mean REAL
@ATTRIBUTE Inst_Loudness_Mean REAL
@ATTRIBUTE A_weighted REAL
@ATTRIBUTE 100_A REAL
@ATTRIBUTE 125_A REAL
@ATTRIBUTE 160_A REAL
@ATTRIBUTE 200_A REAL
@ATTRIBUTE 250_A REAL
@ATTRIBUTE 315_A REAL
@ATTRIBUTE 400_A REAL
@ATTRIBUTE 500_A REAL
@ATTRIBUTE L_1_5 REAL
@ATTRIBUTE L_2 REAL
@ATTRIBUTE L_2_5 REAL
@ATTRIBUTE L_3 REAL
@ATTRIBUTE L_3_5 REAL
@ATTRIBUTE L_4 REAL
@ATTRIBUTE L_4_5 REAL
@ATTRIBUTE L_5 REAL
@ATTRIBUTE L_5_5 REAL
@ATTRIBUTE L_6 REAL
@ATTRIBUTE L_6_5 REAL
@ATTRIBUTE L_7 REAL
@ATTRIBUTE class {Melhor, semelhante, Pior }

@DATA
000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.494,0.494,0.384,0.161,Melhor
000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.044,0.044,0.000,0.000,Melhor
000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.279,0.279,0.399,0.177,Igual
058,0.000,0.000,0.000,0.399,0.399,0.399,0.038,0.000,0.000,0.176,0.176,0.276,0.017,Igual
000,0.000,0.000,0.000,0.520,0.521,0.521,0.120,0.000,0.000,0.284,0.284,0.050,0.000,Melhor
000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.555,0.555,0.663,0.279,Igual
000,0.000,0.000,0.000,0.112,0.113,0.112,0.000,0.000,0.000,0.350,0.351,0.209,0.000,Igual
000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.206,0.206,0.329,0.121,Igual
000,0.000,0.000,0.000,0.315,0.315,0.315,0.043,0.000,0.000,0.564,0.564,0.585,0.099,Pior
174,0.066,0.000,0.000,0.395,0.399,0.399,0.317,0.232,0.229,0.387,0.387,0.491,0.257,Igual
000,0.000,0.000,0.000,0.000,0.003,0.004,0.242,0.251,0.251,0.678,0.678,0.673,0.855,Pior
231,0.044,0.000,0.000,0.053,0.050,0.050,0.169,0.159,0.159,0.598,0.598,0.635,0.868,Pior
000,0.000,0.000,0.000,0.000,0.000,0.000,0.151,0.143,0.143,0.247,0.247,0.354,0.322,Pior
000,0.000,0.000,0.000,0.083,0.084,0.084,0.000,0.000,0.000,0.408,0.408,0.520,0.099,Igual
121,0.170,0.127,0.117,0.107,0.102,0.106,0.917,0.917,0.917,1.000,1.000,1.000,1.000,Pior
000,0.000,0.000,0.000,0.000,0.000,0.000,0.311,0.312,0.312,0.575,0.575,0.633,0.752,Pior
000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.367,0.367,0.424,0.069,Igual
000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.000,0.736,0.736,0.537,0.000,Melhor
000,0.000,0.000,0.000,0.000,0.000,0.000,0.116,0.107,0.107,0.307,0.307,0.420,0.271,Igual

```

Figura 6.1 – Exemplo de formatação ARFF.

Na primeira linha deste formato deve ser indicado o nome do conjunto de dados atribuído pelo comando `@relation nome_do_conjunto_de_dados`. Em seguida, coloca-se a relação de todos os atributos colocando o nome do atributo e tipo utilizando, para isso, o comando `@attribute nome_do_atributo tipo ou {valores}`. Caso os atributos sejam discretos estes devem ser definidos como do tipo real ou inteiro, e caso sejam dados nominais,

indicam-se seus possíveis valores. Para o início da base de dados é usado o comando *@data*, após esse são colocados os exemplos na forma de um por linha com cada valor separado por vírgula.

6.2 - Parâmetros do sistema

Os parâmetros de configuração disponíveis para este algoritmo de RNA são descritos a seguir, de forma a facilitar o entendimento da metodologia utilizada para a busca de uma configuração adequada para a solução do problema de classificação de ruídos.

- Número de neurônio e camadas: descrevem a estrutura básica da rede, sendo necessária a definição do número de camadas escondidas e do número de neurônios de cada camada, pois o número de neurônios de entrada e de saída são definidos a partir do banco de dados.
- Taxa de aprendizado: este fator dita o grau de discretização utilizado para a atualização dos valores dos pesos sinápticos, isto é, altera o gradiente da curva de aprendizado.
- Momento: este termo corresponde a um fator multiplicador que determina um percentual da correção dos pesos da interação anterior que deve ser aplicado à correção atual, fazendo com que as oscilações sejam amortecidas pelo fator de inércia dos pesos assim criado
- Critério de parada de treinamento: é o critério utilizado para interromper o treinamento, podendo ser definido pelo número de iterações, chamadas épocas, ou ainda por um critério de avaliação de erro.

Os conceitos aqui citados são avaliados e otimizados durante a fase de configuração da rede.

6.3 - Metodologia de busca da melhor configuração

Como foi visto anteriormente, na elaboração de uma rede neural artificial pode-se alterar diversas variáveis, dentre elas, o número de neurônios, o número de camadas, a taxa de aprendizado e o momento. Assim, surge o problema de como encontrar a melhor configuração destas variáveis que gere a rede neural com menor erro de predição. Para a solução de problemas desta natureza existem diversos métodos de otimização, mas a maioria não é aplicada para esse caso. Assim, a busca em espaço de estados é uma maneira simples e fácil de ser aplicada a esse tipo de problema, podendo-se, através dela, obter bons resultados.

Como não se conhecia o comportamento da eficiência da rede ao alterar os parâmetros da mesma, elaborou-se uma heurística de busca que consistia em configurar diversas redes neurais seguindo um padrão e observar o erro médio de cada configuração qualificando a melhor, e desta forma realizando uma busca dirigida para se encontrar uma solução boa para o sistema.

Para essa análise de sensibilidade da configuração da rede, inicialmente adotou-se uma configuração padrão de rede neural e a partir desta foram variados os parâmetros da mesma realizando-se uma busca no campo de possíveis soluções. Através de uma análise das redes construídas pelo método *3-fold cross validation* não estratificado (não se mantendo a proporção das classes nos *folds*). Este método consiste na divisão do conjunto de treinamento em três conjuntos de dados (os *folds*), realizando-se o treinamento com dois dos conjuntos e avaliando-se com o terceiro. Dessa forma, são feitos mais dois treinamentos e duas validações permutando-se os conjuntos. O erro final da rede é a média dos erros. Analisou-se a média e

o desvio padrão do erro de cada configuração sempre selecionando em cada análise a melhor configuração e continuando o processo até se obter uma configuração boa.

A seqüência de avaliações apresentadas a seguir descreve o caminho percorrido para a configuração da RNA.

1. Avaliação do critério de parada do treinamento: primeiro fator a ser definido, sendo para tanto, necessário avaliar qual o critério mais adequado: número de épocas ou avaliação de erro.
2. Avaliação do número de neurônios em RNAs de camada única: tendo definido o critério de parada de treinamento, são avaliadas diversas configurações de redes variando-se o número de neurônio em uma rede com uma camada oculta. A análise prosseguirá para a etapa seguinte com algumas das melhores configurações de rede.
3. Avaliação da influência do número de camadas: variando-se o número de camadas e o número de neurônio em cada uma delas percorre-se um espaço de soluções com o objetivo de selecionar as melhores configurações. Esta análise pode ser dividida em redes balanceadas, onde, definido o número de neurônios de uma camada oculta, distribui-se este número pelo número total de camadas ocultas, e desbalanceadas, onde, definido o número de neurônios de uma camada oculta, repete-se este número para as demais camadas.
4. Avaliação do momento: estando definido o número de camadas e o número de neurônios de cada uma avalia-se então o valor do momento da rede. Este parâmetro diminui a possibilidade de que a rede estabilize em mínimos locais.

5. Avaliação da taxa de aprendizado: é a última análise de otimização da rede. Com este parâmetro procura-se otimizar o treinamento da rede, fazendo com que ela atinja de forma mais rápida o ponto de treinamento desejado.

6.4 - Configuração do classificador considerando os atributos dB(A) e loudness

A configuração desta RNA foi baseada no banco de dados citado no Capítulo 5, contendo as informações subjetivas do júri, os valores em terço de oitava do espectro em dB(A) e os valores de *loudness* globais e em escala de banda crítica.

6.4.1 - Configuração inicial

Para essa análise, inicialmente, adotou-se uma configuração padrão de rede neural artificial e a partir de então foram montadas diversas configurações variando-se os parâmetros da mesma, realizando assim uma busca em espaço de estados.

- Número de camadas: 1
- Número de neurônios: $(\text{atributos} + \text{classes}) / 2 = 35$
- Taxa de aprendizado: 0,3
- Momento: 0,2
- Épocas: 500

A configuração descrita acima, foi utilizada também para a avaliação do critério de parada do treinamento.

6.4.2 - Critério de parada do treinamento

Este parâmetro dita em que momento o treinamento deve ser interrompido, seja por número de ciclos, normalmente denominado por número de épocas, ou ainda por uma avaliação do percentual de erro de uma amostra dos dados de validação.

Para a avaliação do número de épocas de aprendizado fez-se uma variação de 100 à 1000 em intervalos de 50 épocas, de 1000 à 5000, em intervalos de 500 épocas. O resultado obtido é apresentado na Figura 6.2.



Figura 6.2 – Avaliação do número de épocas de treinamento.

Observa-se a insensibilidade deste parâmetro como critério de parada, sendo difícil avaliar o momento adequado para interromper o treinamento. É válido observar que uma busca mais refinada na região que apresentou os menores erros poderia ser realizada. Contudo, sabe-se que estes valores variam de uma configuração de rede para outra, necessitando-se de um critério mais adequado.

A outra maneira de interromper o treinamento, que é disponibilizada no software Weka, é determinar o percentual do conjunto de validação que será utilizado para a verificação do erro. Assim, quando o erro deste conjunto passar a aumentar, uma outra variável passará a contar o número de ocorrências deste aumento, isto é, determina o número de vezes que o erro pode aumentar até a parada do treinamento. O valor padrão é 20.

A Figura 6.3 apresenta a relação entre o erro da rede e o tamanho do conjunto deste conjunto.

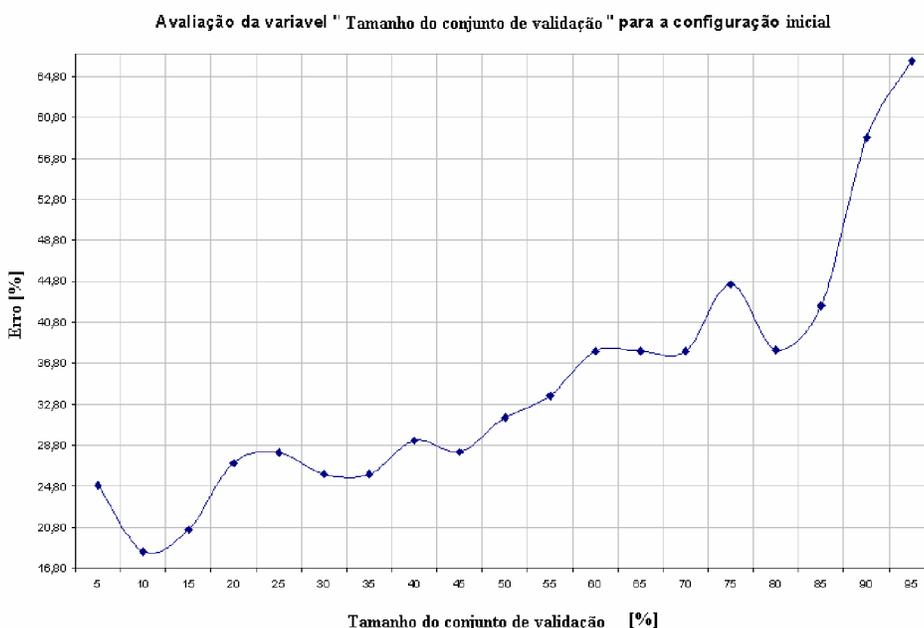


Figura 6.3 – Avaliação do percentual do conjunto de validação a ser utilizado para interromper o treinamento.

Diferentemente do critério anterior, o número de épocas, a avaliação do erro por este critério apresenta uma boa sensibilidade.

Definido o tamanho percentual do conjunto de validação, as análises continuarão de maneira que a variável que determina o número de ocorrências de aumento de erro, o “*validation threshold*”, continuará sendo variado, isso para que verifique a tendência das curvas de erro.

Com o critério de parada definido, iniciam-se as avaliações de configuração da RNA de classificação com um erro da rede padrão de aproximadamente 18%.

6.4.3 - Número de neurônios na camada oculta

Para verificar a influência do número de neurônios na primeira camada, analisaram-se algumas redes alterando o valor de neurônios de 10 a 40 em intervalos de 5 em 5 neurônios. Os resultados obtidos podem ser observados na Figura 6.4.

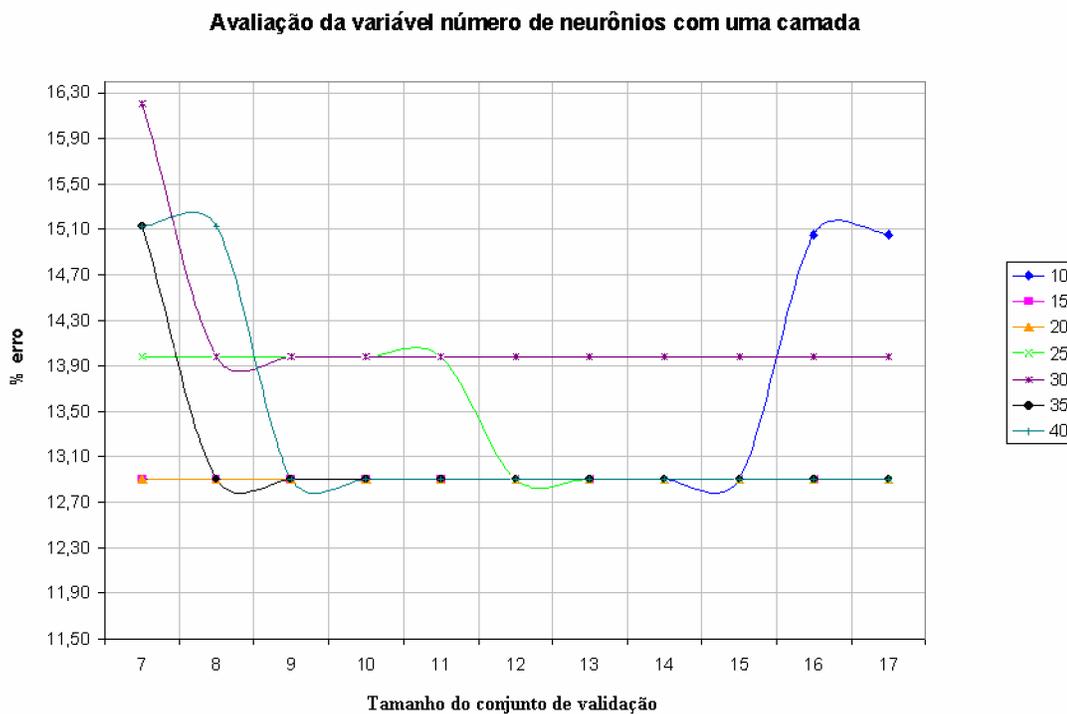


Figura 6.4 – Avaliação do número de neurônios na camada oculta.

Verificando-se uma tendência de diminuição do erro da rede para valores entre 30 e 40 neurônios, em busca de melhores resultados, realizou-se uma variação mais refinada de 1 em 1 neurônio no intervalo de 30 a 40 neurônios, sendo os resultados apresentados na Figura 6.5.

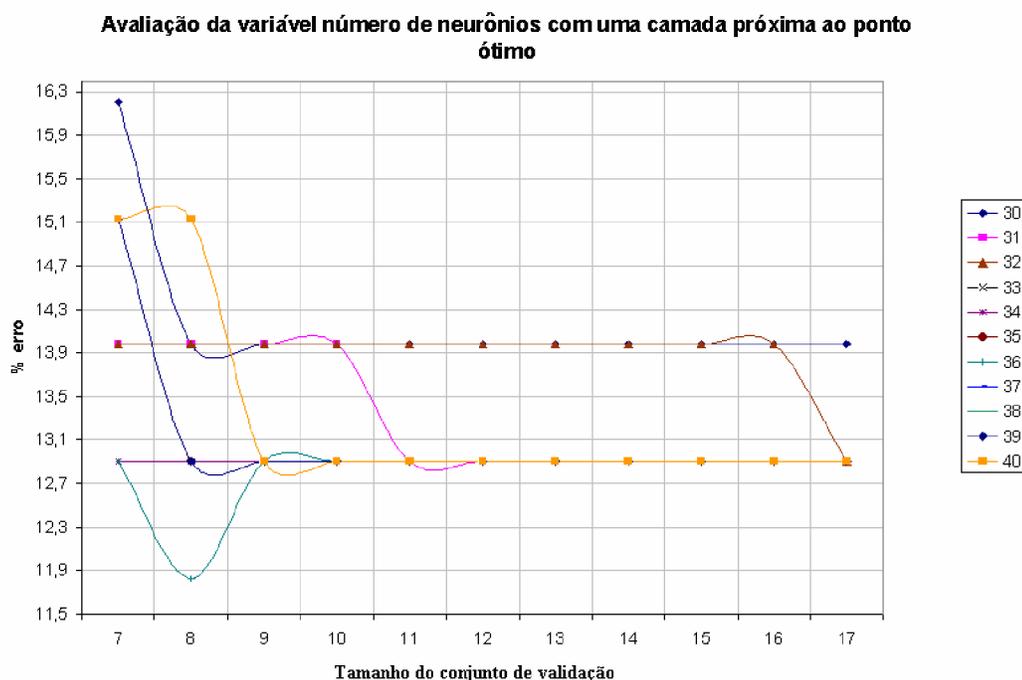


Figura 6.5 – Avaliação do número de neurônios camada oculta próximo ao ponto ótimo.

Pode-se observar que o melhor resultado foi apresentado pela configuração com 36 neurônios na camada oculta. Esta, uma melhora significativa, pois o erro médio da rede passou de 18% para cerca de 12%.

6.4.4 - Número de camadas

Em uma próxima análise deve-se avaliar a influência do número de camadas. Neste ponto podem-se adotar dois critérios, sendo o primeiro adicionando camadas desbalanceadas e o segundo, camadas balanceadas.

a) Camadas Desbalanceadas:

Nesta análise variou-se o número de camadas em 1, 2 e 3 camadas, mantendo-se em cada camada o mesmo número de neurônios.

Avaliação da variável " Tamanho do conjunto de validação " - Variação de camadas

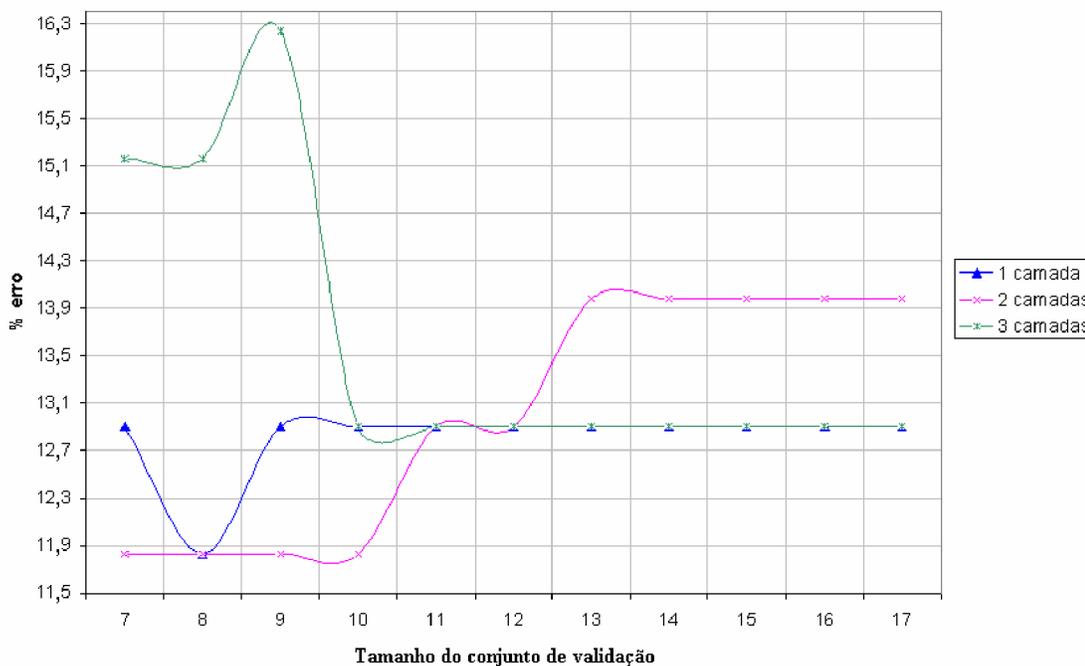


Figura 6.6 – Avaliação do número de camada desbalanceadas em 36 neurônios.

b) Camadas Balanceadas

Como comparação foi realizada uma avaliação variando-se o número de camada e mantendo-se a proporção no número de neurônios da rede. Assim, foram configuradas as RNA com 1, 2 e 3 camadas, onde cada camada possuía a seguinte configuração balanceada em 36 neurônios

1 camada–36 neurônios

2 camadas–18,18 neurônios

3 camadas–12,12,12 neurônios

A Figura 6.8 apresenta os resultados desta validação.

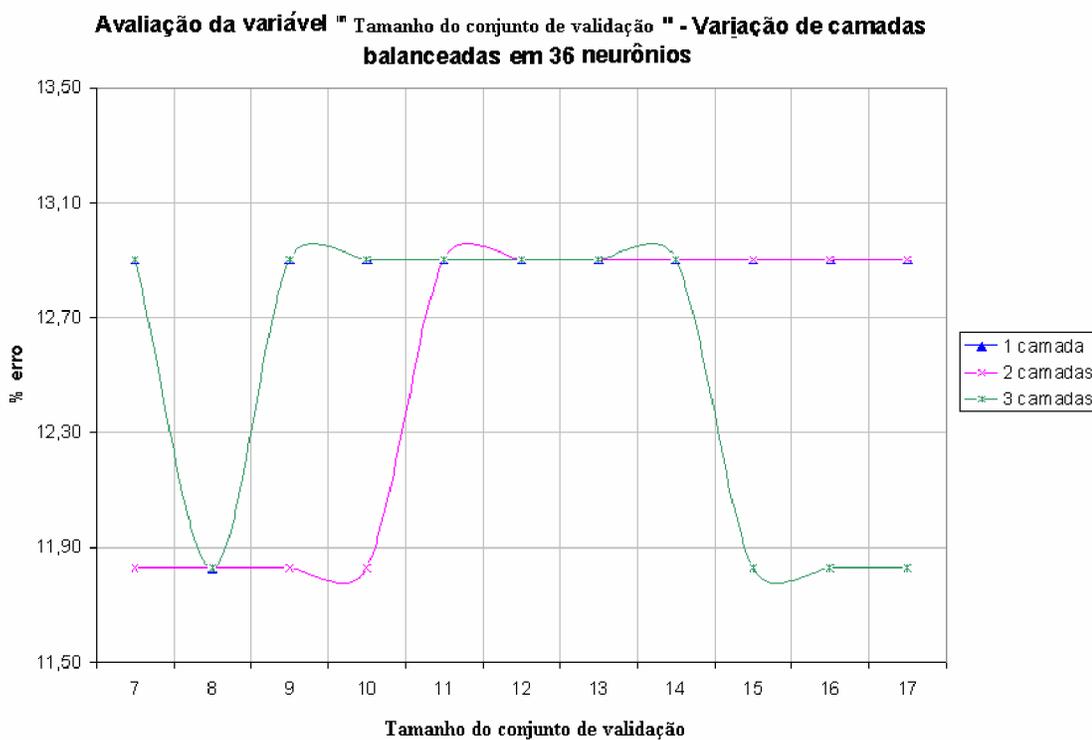


Figura 6.7 – Avaliação do número de camada balanceadas em 36 neurônios.

Tendo diferentes configurações apresentado resultados semelhantes, optou-se por continuar as análises com as 3 das melhores configurações, continuando assim com as seguintes redes:

- RNA com 1 camada com 36 neurônios
- RNA com 2 camadas com 18 neurônios por camada
- RNA com 3 camadas com 12 neurônios por camada

Observa-se, que com estas modificações não foi possível incrementar o desempenho da rede, restando para tanto a avaliação dos parâmetros de momento e do coeficiente de aprendizado.

6.4.5 - Momento

Da mesma forma que foram avaliados os parâmetros anteriores, variou-se o momento de forma discreta, em intervalos de 0,1 para as 3 melhores configurações de redes neurais. A variável momento pode variar no intervalo de (0:1). Os resultados são apresentados na Figura 6.8.

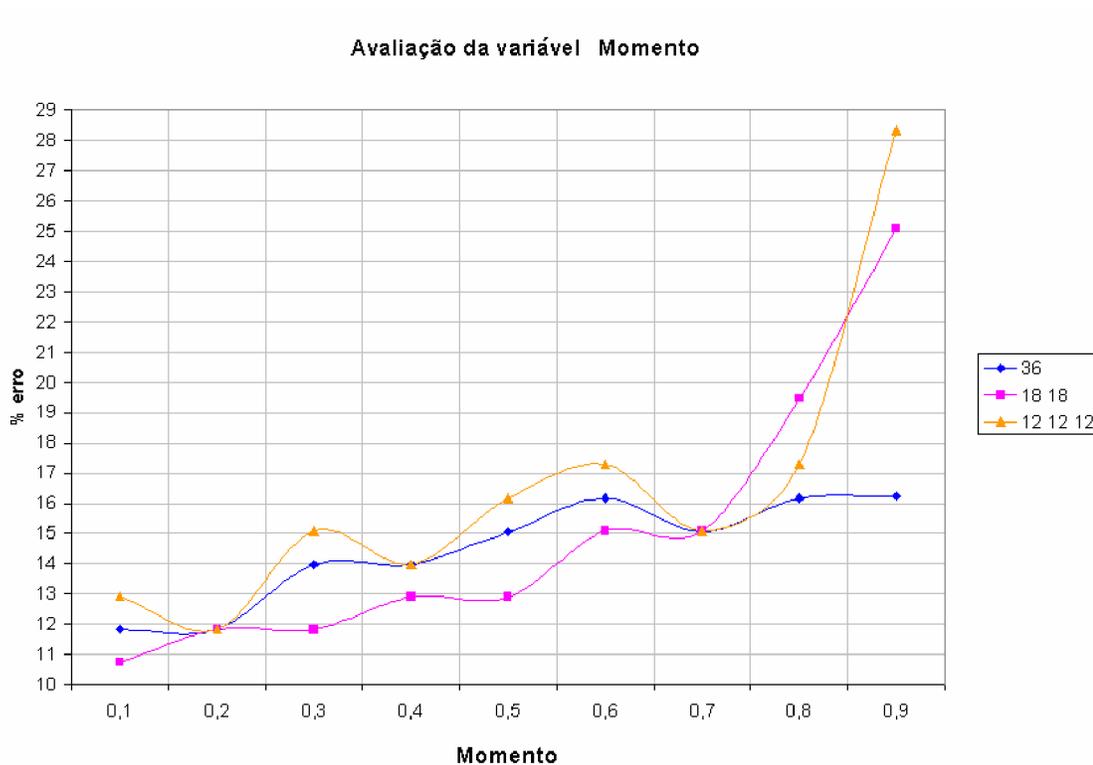


Figura 6.8 – Avaliação do parâmetro momento.

A configuração de duas camadas com 18 neurônios em cada apresentou o melhor resultado, chegando a um erro médio da ordem de 10%. Contudo, para a próxima análise, a taxa de aprendizado, serão mantidas as três configurações acima analisadas. Para a próxima análise são adotados o valor de momento com melhores resultados, ou seja, igual a 0,1.

6.4.6 - Taxa de aprendizado

A última análise a ser realizada é a da variável denominada taxa de aprendizado. Esta tem o mesmo domínio da variável anterior, ou seja, deve estar no intervalo de (0:1). Na Figura 6.9 estão os resultados desta análise.

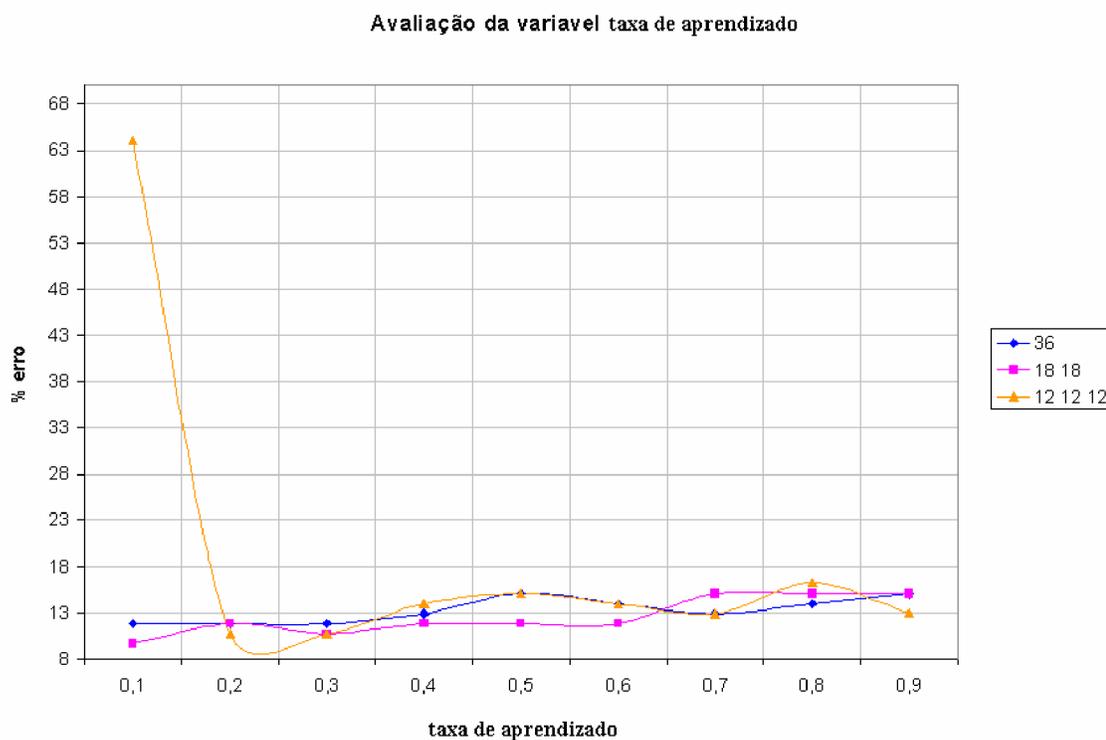


Figura 6.9 – Avaliação do parâmetro taxa de aprendizado.

Com estes resultados finaliza-se a análise de sensibilidade dos parâmetros que caracterizam e configuram a rede neural. O erro apresentado por esta configuração, cerca 8,7%, é considerado um valor aceitável, tendo em vista o número de exemplos disponíveis para o treinamento desta RNA. Vale lembrar que cada avaliação errônea equivale a pouco mais de três por cento de erro.

6.4.7 - Configuração final

Tendo finalizado esta etapa da configuração do classificador, são listados abaixo os parâmetros da rede que apresentou o menor erro médio dentre as configurações avaliadas:

- 2 camadas com 18 neurônios e cada camada;
- tamanho do conjunto de validação: 10;
- variável de validação *threshold*: 8;
- momento: 0,1;
- coeficiente de aprendizado: 0,1.

Taxa de erro encontrada com as configurações acima: 8,7 %.

Uma última análise da rede neural aqui configurada, permitiu identificar uma característica interessante das RNAs. Esta característica baseia-se no postulado de aprendizado de Hebb [14], que avalia a utilização das sinapses dos neurônios. Este, viabiliza a utilização da RNA não só como classificador, mas também como identificador de atributos relevantes para o sistema, isto é, os neurônios dos atributos (métricas) que tiveram suas sinapses enfraquecidas ou eliminadas, provaram ser valores desnecessários ou inconsistentes para a correlação dos padrões identificados pela rede. Dessa forma uma nova configuração de rede será avaliada com o objetivo de selecionar os valores de entrada, ou seja as métricas, que são significativas para o processo de classificação da RNA.

6.5 - Configuração do classificador considerando os atributos dB(A), loudness, sharpness, roughness e fluctuation strenght

Conforme descrito no item anterior, o objetivo desta nova configuração de RNA consiste não apenas na elaboração de um classificador, mas também na identificação dos atributos importantes para a identificação dos padrões dos exemplos utilizados, no caso em questão as métricas. Dessa forma as saídas desta avaliação é uma nova configuração do classificador e a identificação das métricas que melhor correlacionam os dados de ruído com a avaliação do júri.

O banco de dados para esta nova avaliação contém os valores de *sharpness*, *roughness* e *fluctuation strenght*, em escala de banda crítica, e os valores de *proeminence ratio*. A matriz de entrada passa então a conter cento e cinquenta e três linhas e noventa e duas colunas. Também fazendo-se necessário a transposição da mesma.

A metodologia utilizada é a mesma do desenvolvimento da rede anterior. Assim os tópicos serão apresentados de forma resumida.

6.5.1 - Configuração inicial

De forma semelhante à análise anterior os parâmetros iniciais são os valores padrões, agora com a diferença do número de neurônios de entrada e da camada oculta, tendo respectivamente cento e cinquenta e dois e setenta e seis.

6.5.2 - Critério de parada do treinamento

Com o mesmo comportamento da RNA anterior, o critério de parada também foi o de avaliação de erro.

6.5.3 - Número de neurônios na camada oculta

Novamente é realizada a avaliação do número de neurônios

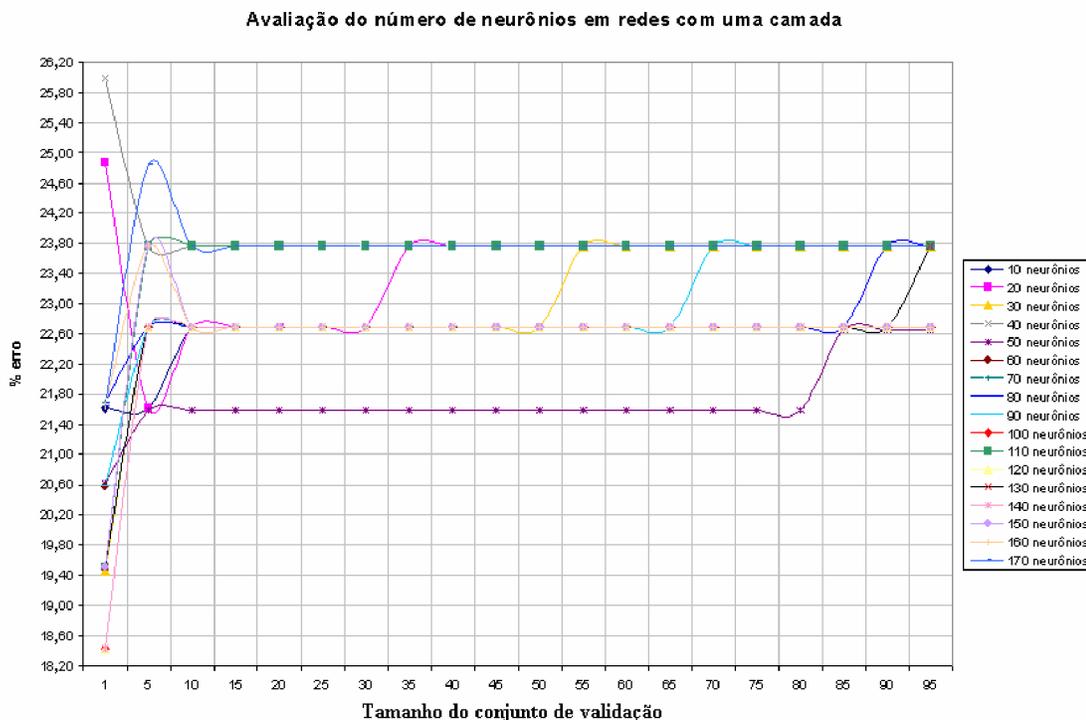


Figura 6.10 – Avaliação do número de neurônios na camada oculta.

Observa-se a média de erros inicia-se em valores próximos a da avaliação anterior, tendo também um mínimo em torno dos 18%.

A Figura 6.11 apresenta a avaliação da tendência do erro médio para o critério de parada de treinamento *validation treshold* igual a um, para verificar os resultados apresentados na Figura 6.10.

Avaliação da tendência do erro médio para Validation Threshold igual a 1

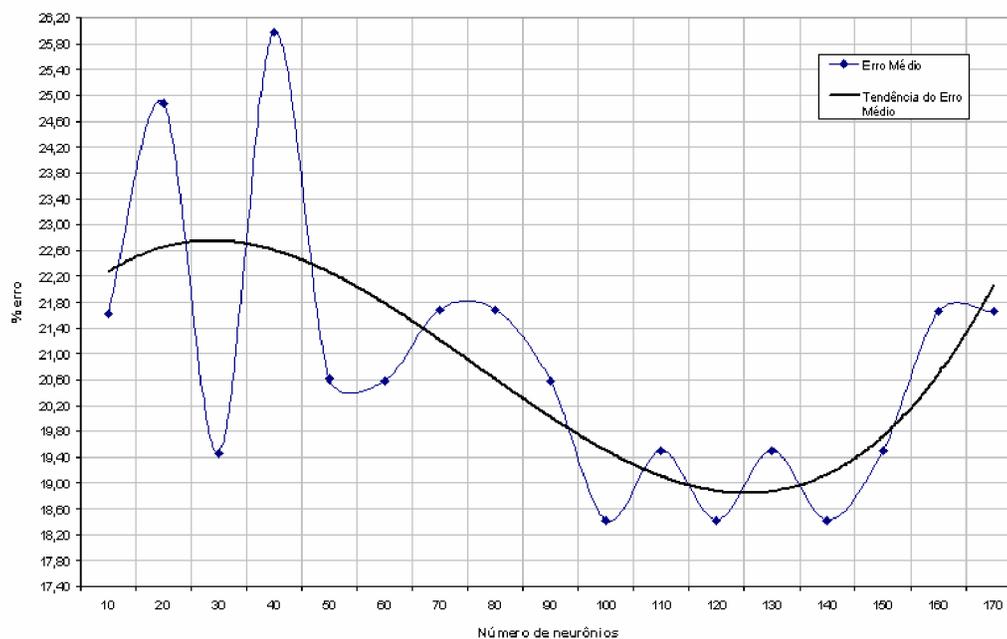


Figura 6.11 – Avaliação da tendência do erro médio para a variável *validation treshold* igual a um.

Da mesma forma que os resultados anteriores o ponto mínimo do intervalo está próximo dos cento e trinta neurônios na camada escondida. Mais uma análise foi realizada a fim de melhorar o desempenho da RNA. Contudo, o menor erro médio encontrado foi mesmo da análise acima.

6.5.4 - Número de camadas

A avaliação do número de camadas ocultas foi realizada de forma semelhante às análises anteriores. Os resultados são apresentados na Figuras 6.13 e 6.14.

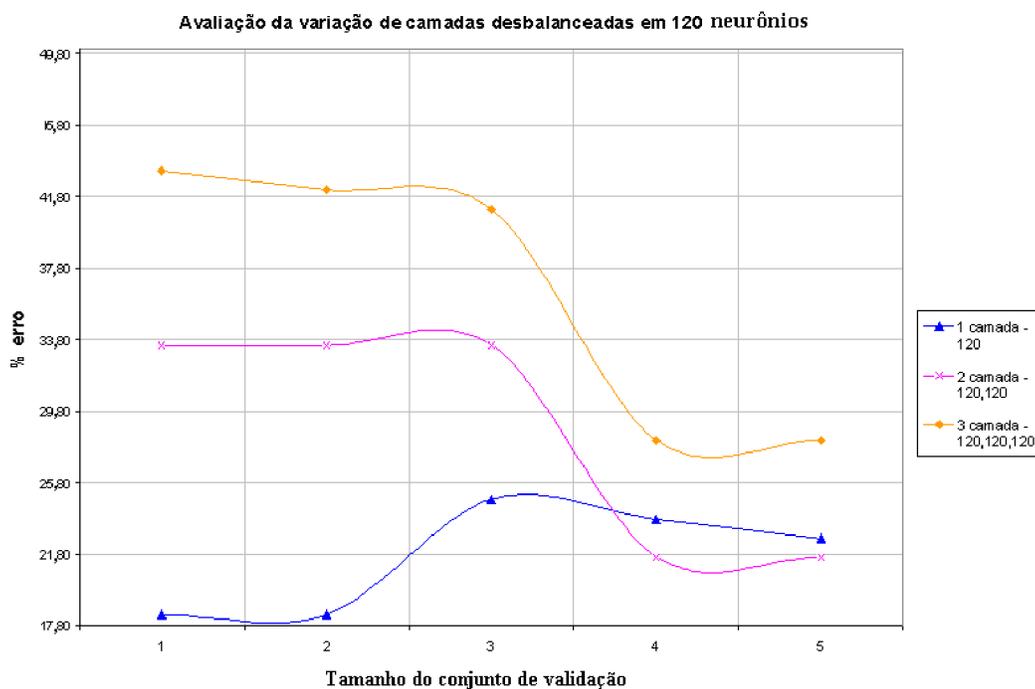


Figura 6.12 – Avaliação do número camadas desbalanceadas em 120 neurônios.

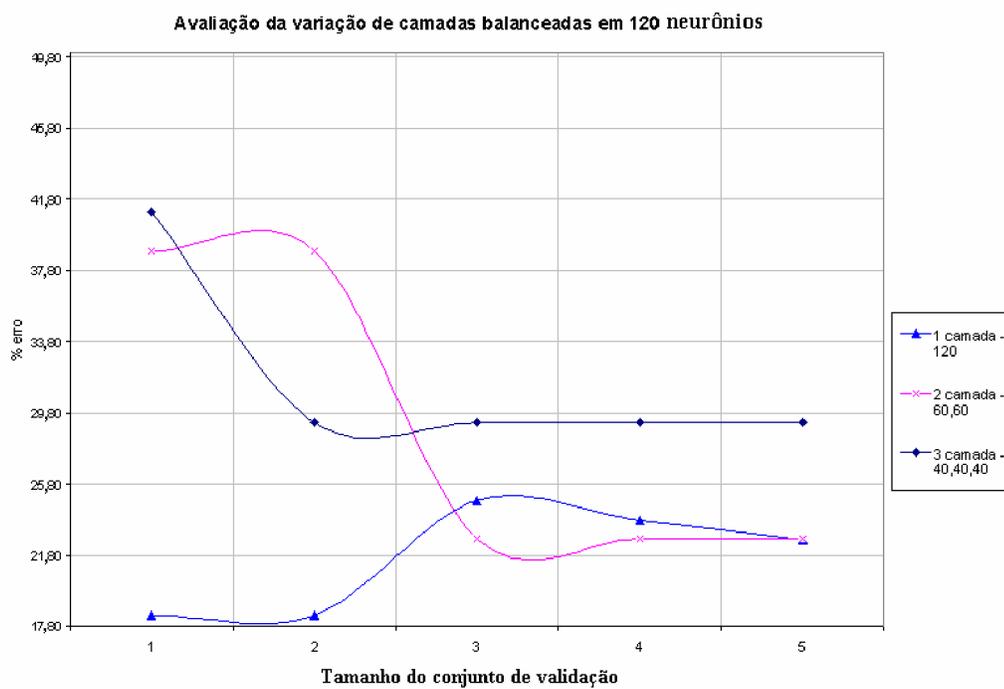


Figura 6.13 – Avaliação do número camadas balanceadas em 120 neurônios.

Diferentemente da situação da outra RNA não houve melhoras no erro médio da rede, permanecendo em 18%.

6.5.5 - Momento

A avaliação da variável momento também não acrescentou melhoras para a rede.

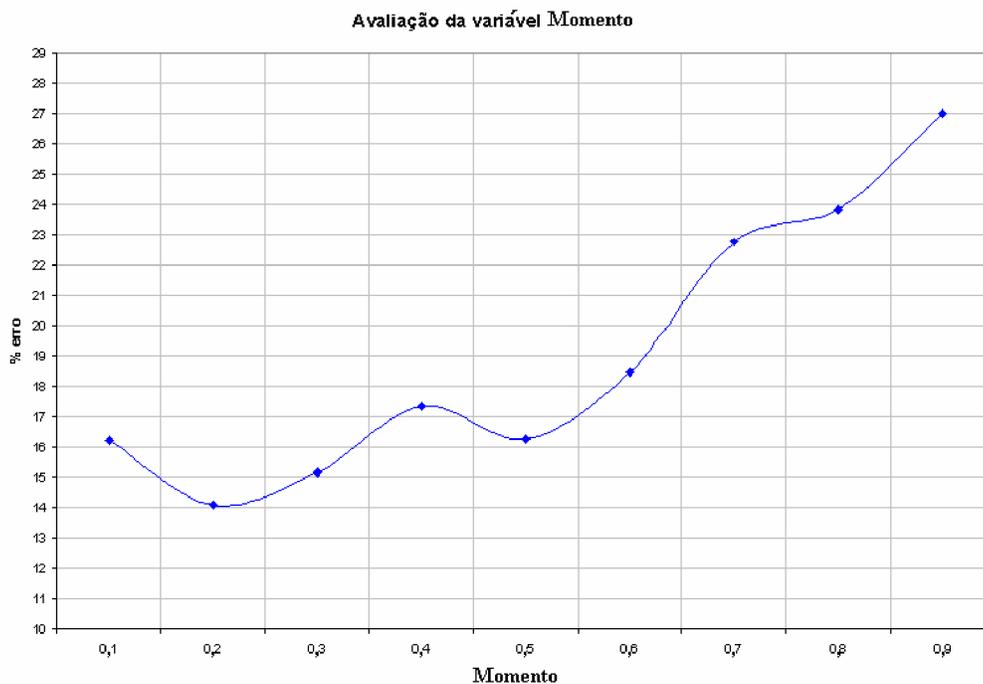


Figura 6.14 – Avaliação do parâmetro de momento.

6.5.6 - Taxa de aprendizado

A taxa de aprendizado é o último parâmetro deste conjunto de análises. Também efetuada conforme a metodologia anterior, variou-se o seu valor com o objetivo de encontrar o melhor resultado da faixa de avaliação. Novamente não houve incremento no percentual de acertividade da RNA. A avaliação da taxa de aprendizado é apresentada na Figura 6.15.

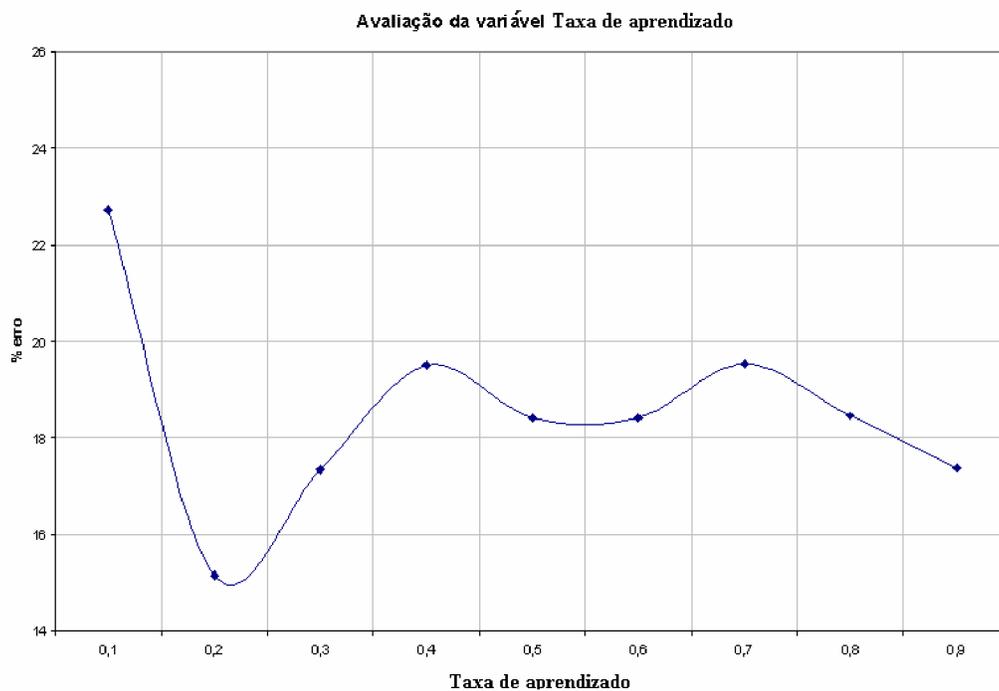


Figura 6.15 – Avaliação do parâmetro taxa de aprendizado.

6.5.7 - Configuração final

Assim, a configuração final do classificador dB (A) / *loudness* / *sharpness* / *roughness* / *fluctuation strenght* é apresentada abaixo:

- 1 camada com 120 neurônios;
- Tamanho do conjunto de validação: 2;
- Variável de validação *threshold*: 2;
- Momento: 0,2;
- Coeficiente de aprendizado: 0,2.

Taxa de erro encontrada com as configurações acima: 14 %.

6.5.8 - Avaliação geral das performances das redes neurais artificiais

É interessante observar que o resultado apresentado pela segunda configuração de RNA foi um pouco pior do que a primeira, tendo respectivamente 14% e 9% de erro médio aproximado. Para a verificação da eficiência destas duas configurações um último teste foi realizado. Sabendo-se que a ordem do conjunto de treinamentos tem influência no desempenho das RNAs, o conjunto de banco de dados para o treinamento das redes foi estratificado, ou seja, distribuiu-se o conjunto de dados em três conjuntos, de maneira que se mantivessem as proporções das classes em cada conjunto. Novos ciclos de treinamento foram realizados. Os erros médios de treinamentos dos dois classificadores diminuiram, sendo que seus erros finais estão apresentados na Tabela 6.1.

Tabela 6.1 – Erro médio das configurações de RNAs finais

Classificador	Erro médio aproximado
<i>dB(A) / loudness</i>	6%
<i>dB(A) / loudnes / sharpness / roughness / fluctuation strength</i>	8%

Assim pode-se afirmar que as RNAs têm desempenhos parecidos. Avaliando-se ainda o desempenho do segundo classificador conclui-se que vários caminhos podem ser utilizados para o processo de configuração de RNAs, e que os atributos que não são determinantes para a geração de conhecimento das RNAs podem ser identificados e eliminados da base de dados, sem que esta diminua seu desempenho.

Raciocinando-se sobre os erros apresentados pelas configurações encontradas, pode-se separá-los em duas parcelas. Uma referente a estrutura e configuração dos parâmetros da

rede e outra referente às incertezas do dados da análise de júri (fase de aprendizado). Estes podem ser minimizados aumentando-se o número de amostras do conjunto de treinamento e aprimorando-se o tratamento estatístico dos dados do júri. Pela literatura, o conjunto de dados utilizado neste trabalho é considerado extremamente pequeno, casos típicos apresentados mostram conjuntos da ordem de mil amostras. Contudo, a dificuldade de avaliação de júri inviabiliza valores desta ordem de grandeza.

CAPÍTULO 7 - CONCLUSÕES

A realização deste trabalho envolveu basicamente duas ferramentas distintas para a avaliação da qualidade sonora de compressores herméticos, a análise de júri e a rede neural artificial.

O processo de preparação e realização da análise de júri, apresentado neste trabalho, revelou inúmeros cuidados especiais a serem tomados para sua aplicação. Um exemplo claro destes cuidados refere-se à ordem de amostragem dos ruídos a serem avaliados, na qual, identificou-se uma tendência dos jurados avaliarem ruídos, outrora considerados semelhantes à referência, como sendo ruídos mais agradáveis do que a referência quando estes sucediam um ruído de nível elevado. Exemplos como este, somados a outros fatores do processo, e aliados ainda à dispersão de opiniões dos jurados, vêm confirmar a necessidade de estudos mais profundos e específicos, no que diz respeito à aplicação de análise de júri para avaliação da qualidade sonora.

O objetivo principal deste trabalho, demonstrar a possibilidade de utilização de RNAs para avaliação da qualidade sonora, foi alcançado com a utilização de um modelo simples de rede, o *perceptron* de múltiplas camadas. A grande dificuldade desta aplicação está na formação do banco de dados, que por estar diretamente ligado ao processo de avaliação subjetiva e necessitar de um grande número de amostras, torna o processo de elaboração do banco de dados um trabalho custoso. A aplicação das redes neurais artificiais como ferramenta de classificação no processo de avaliação da qualidade sonora, demonstrou grande potencial, tendo apresentado bons resultados quando levada em consideração a quantidade de exemplos utilizada nos treinamentos.

Uma característica importante das regras de aprendizado das redes neurais, baseada no postulado de aprendizado de Hebb [14] “Se dois neurônios em ambos os lados de uma sinapse são ativados assincronamente, então aquela sinapse é seletivamente enfraquecida ou

eliminada” possibilitou a utilização da RNA não só como classificador, mas também como identificador de atributos relevantes para o sistema, isto é, os neurônios dos atributos (métricas) que tiveram suas sinapses enfraquecidas ou eliminadas, provaram ser valores desnecessários ou inconsistentes para a correlação dos padrões identificados pela rede. Tal característica permitiu a identificação dos valores de *loudness* e dB (A) como sendo representativos das avaliações subjetivas apresentadas na análise de júri. O trabalho apresentado por Nunes [23], compartilha desta conclusão.

Sugestão para trabalhos futuros

Devido ao potencial apresentado pela utilização das redes neurais como forma de substituir um conjunto de jurados para a avaliação da qualidade sonora de uma determinada classe de produtos, incentiva-se a utilização e experimentação de outros modelos de neurônios capazes de gerarem aprendizado de forma não supervisionada, de modo a possibilitar a utilização de um banco de dados suficientemente grande, reduzindo-se a quantidade de ruídos a serem amostrados para o júri, uma vez apenas algumas amostras de cada grupo de ruídos será suficiente para a identificação dos grupos de ruído.

Outro modelo de rede neural que pode ser aplicado é o modelo de redes recorrentes. Com um modelo de redes recorrentes seria possível a utilização de sinais de ruído no domínio do tempo, possibilitando agregar informações e características apresentadas por este sinal. Em suma, esta ferramenta deve ser explorada de forma mais profunda, haja vista seu potencial de aplicação em problemas desta natureza.

Devido à importância da qualificação dos sinais, formas mais confiáveis de realização do júri, considerando características psicológicas devem, ser trabalhadas e estudadas.

REFERÊNCIAS BIBLIOGRÁFICAS

1. LYON, Richard H. Designing for product sound quality. New York. Marcel Dekker, Inc.: 2000.
2. NUNES, Rafael A. S. Avaliação do ruído de compressores herméticos e de sistemas de refrigeração através da qualidade sonora. Dissertação de mestrado, UFSC 2001.
3. ARRUDA, J. R. F. Jornal da unicamp. Campinas: setembro de 2003
4. NEPOMUCENO, L. A. Elementos de acústica física e psicoacústica. São Paulo, Edgard Blücher LTDA: 1994.
5. FASTL H., ZWICKER E. Psychoacoustics, facts and models. Berlin. Springer. 2 ed.: 1999.
6. BRÜEL & KJAER. Sound quality workshop. Norcross: 1999.
7. LENZI A. Apostila de segurança do trabalho. Florianópolis: 2001
8. ISO. Acoustics – Method for calculating loudness level – ISO 532. 1975.
9. KRYTER, Karl D. The effects of noise on man. New York. Academic Press: 1970.
10. LIMA, F. R; GERGES S. N. Y.. Qualidade Sonora para diagnóstico e controle de qualidade veicular. Seminário de acústica – AEA: 2004.
11. FASTL H., Psychoacoustics and sound quality metrics. Michigan. Sound quality symposium.: 1998.
12. NEUTRIX CORTEX INSTRUMENTS. Psicoacústica – Uma ferramenta para projetos industriais de áudio. Santo Angelo, Guarulhos: 1999.
13. LYON, R. H. Product Sound Quality – From Percetion to Design. Massachusetts, 2003.

14. HAYKIN, S. Redes Neurais – princípios e práticas. Porto Alegre, Bookman, 2001 segunda edição.
15. FU, L. Neural networks in computer intelligence. New York, MacGraw Hill, 1994
16. DENCKER F. A. Detecção de falha na montagem de compressors herméticos por redes neurais artificiais. Dissertação de mestrado, UFSC 2002.
17. McCULLOCH. W. S., PITTS, W. A logical calculus of ideas immanent in nervous activity. Bulletin of Mathematical Biophysics. N. 5 p.115-133, 1943.
18. FREEMAN, J. A.; SAKAPURA, D. M. Neural Networks: Algorithms, Applications, and Programming Techniques. Addison-Wesley, 1991.
- 19 HERTZ, Addison-Wesley, 1991. Z, J.; et. Al. Introduction to the theory of neural computation. Redwood City, CA
21. WIDROW. B.; LEHR, M. A. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. Proceedings of the IEEE, v.78, n.9, p.1415-1441, Sept. 1990.
22. RUMELHART, D. E.; McCLELLAND, J. L. Parallel distributed processing. Sna Diego, MIT Press, 1986.
23. NUNES, Rafael A. S., LENZI, A., BAARS, E. Qualidade sonora de compressores herméticos e refrigeradores. COBEM: 2002
24. BRAGA, A. P.; LUDEMIR, T., B.; CARVALHO, A.C. Redes neurais artificiais – Teoria e aplicações. Rio de Janeiro. LTC, 2000.
25. STEINBERG, B. Z.; et. al. A neural network approach to source localization. Washington, 1991.

26. KAMINSKI, P. C.; MAGALHÃES F. S. Utilização de redes neurais artificiais e frequências naturais para a detecção de múltiplos defeitos. Congresso ibero latino americano sobre métodos computacionais para engenharia, 1995.
27. SUSINI, P.; et. al. Characterizing the sound quality of air-conditioning noise. Paris: 2004.
28. BODDEN M.; et. al. Sound quality evaluation of interior vehicle noise using an efficient psychoacoustic method. Alemanha. Euronoise, 1998.