

**PAULO RICARDO FRANCHI ZULATO**

**ANÁLISE DE PADRÕES DE  
VOCALIZAÇÃO PARA DETERMINAR  
O BEM-ESTAR DE SUÍNOS**

Trabalho de Conclusão de Curso apresentado  
à Escola de Engenharia de São Carlos, da  
Universidade de São Paulo

Curso de Engenharia Elétrica com ênfase em  
Eletrônica

ORIENTADOR: Prof. Dr. Rodrigo Capobianco Guido

São Carlos

2010

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica preparada pela Seção de Tratamento  
da Informação do Serviço de Biblioteca – EESC/USP

Z94a	<p>Zulato, Paulo Ricardo Franchi</p> <p>Análise de padrões de vocalização pra determinar o bem-estar de suínos / Paulo Ricardo Franchi Zulato ; orientador Rodrigo Capobianco Guido. -- São Carlos, 2010.</p> <p>Trabalho de Conclusão de Curso (Graduação em Engenharia Elétrica com ênfase em Eletrônica) -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2010.</p> <p>1. Processamento digital de sinais. 2. Suínos. 3. Redes neurais artificiais. 4. Análise de ondaletas. I. Título.</p>
------	---

A meus pais, que abdicaram de muitos de seus planos pessoais para que eu e minha irmã pudéssemos realizar os nossos.

## **Agradecimentos**

Agradeço primeiramente a Deus, pela vida, por me guiar às escolhas certas e pela chance de concluir mais uma etapa em minha vida.

Aos meus pais, Antônio e Beatriz, que sempre me acompanharam e apoiaram em todas as fases, sempre com muita compreensão e carinho.

À minha irmã, Cecília, pela companhia e por sempre ter confiado em mim e me incentivado a continuar meu caminho.

À minha namorada, Aline, por estar sempre ao meu lado, pelo carinho e por me suportar em meus dias ruins.

Ao meu orientador, Rodrigo Capobianco Guido, por ter aceitado me orientar neste trabalho, pela paciência, confiança e tempo dedicado em meu auxílio.

À Késia e à Natalia, por terem confiado a mim a realização do projeto por elas planejado.

A todos os meus amigos que de alguma forma me ajudaram em todos esses anos.

A todos aqueles que contribuíram para meu crescimento pessoal e profissional durante este período.

# Sumário

<b>1.</b>	<b>Introdução .....</b>	<b>1</b>
<b>2.</b>	<b>Objetivos .....</b>	<b>1</b>
<b>3.</b>	<b>Conceitos Iniciais .....</b>	<b>2</b>
3.1.	Bem-estar de Suínos .....	2
3.2.	Digitalização de Sinais .....	4
3.3.	Filtros .....	6
3.4.	Redes Neurais Artificiais .....	8
3.5.	Transformada <i>Wavelet</i> .....	18
3.6.	Bandas Críticas .....	25
3.7.	Dimensão Fractal .....	26
<b>4.</b>	<b>Descrição das Atividades .....</b>	<b>28</b>
<b>5.</b>	<b>Resultados .....</b>	<b>32</b>
<b>6.</b>	<b>Conclusão .....</b>	<b>38</b>
<b>7.</b>	<b>Referências Bibliográficas .....</b>	<b>39</b>

## Lista de Figuras

Figura 3.2.1 - Amostragem de um sinal sonoro qualquer.....	5
Figura 3.2.2 - Discretização do som.....	6
Figura 3.3.1 - Parâmetros de um filtro (Júnior, 2007).....	7
Figura 3.4.1 - Neurônio simplificado.....	9
Figura 3.4.2 - Modelo de um neurônio artificial .....	10
Figura 3.4.3 - Função Gaussiana.....	11
Figura 3.4.4 - Esquema de uma Rede Neural Artificial.....	12
Figura 3.4.5 - Função de base radial do tipo gaussiana.....	14
Figura 3.4.6 - Fronteira de separabilidade RBF .....	14
Figura 3.5.1 - Comparação entre as análises (a) de Gabor (STFT) e (b) Multirresolucional (Oliveira, 2007).....	20
Figura 3.5.2 - Comparação da análise espectral entre (a) STFT e (b) WT (Oliveira, 2007).....	22
Figura 3.5.3 - Funcionamento de uma DWT (Junior, 2007).....	24
Figura 4.1 - Articulação de leitões acometidos por artrite.....	29
Figura 4.2 - Gravador Digital Diretivo Panasonic®RR-US395.....	29
Figura 4.3 - Separação de cada grito .....	30
Figura 4.4 - Metodologia de aquisição e análise dos sinais de áudio .....	32
Figura 5.1 - Resultado utilizando o primeiro grupo de maior entalpia.....	34
Figura 5.2 - Resultado utilizando o segundo grupo de maior entalpia .....	35
Figura 5.3 - Resultado utilizando o terceiro grupo de maior entalpia.....	35
Figura 5.4 - União dos resultados dos teste de entalpia.....	36
Figura 5.5 - Resultado utilizando o primeiro grupo de maior ITGU.....	37
Figura 5.6 - Resultado utilizando o segundo grupo de maior ITGU .....	37
Figura 5.7 - Resultado utilizando o terceiro grupo de maior ITGU.....	37
Figura 5.8 - União dos resultados dos testes de ITGU.....	38

## Lista de Abreviaturas e Siglas

AMR	Análise Multirresolucional
CAD	Conversão Analógica-Digital
CWT	<i>Continuous Wavelet Transform</i>
DWT	<i>Discrete Wavelet Transform</i>
ITGU	Índice de Temperatura de Globo e Umidade
RBF	<i>Radial Basis Function</i>
RNA	Redes Neurais Artificiais
STFT	<i>Short Time Fourier Transform</i>

## Resumo

O objetivo deste trabalho é propor um algoritmo para realizar a identificação de padrões na vocalização suína, visando determinar o nível do bem-estar do animal. Tal análise foi proposta uma vez que o bem-estar animal é um assunto cada vez mais abordado no mundo todo, principalmente quando os animais são criados para o abate. Dessa forma, a criação de um método em que haja o mínimo de contato com os animais se faz importante, evitando que tal contato altere o comportamento do animal e, conseqüentemente, o resultado da análise de seu bem-estar. Por essas características, foi proposto um método de análise dos sons emitidos pelos suínos com base na utilização de uma Rede Neural Artificial do tipo *Radial Basis Function*, a qual possui como elementos de treinamento e operação um conjunto de características extraídas através da Transformada Discreta *Wavelet* de sinais sonoros pré-gravados. As características obtidas dos sinais foram as energias das bandas críticas relativas à Escala Bark e a diferença entre as energias das bandas adjacentes, além dimensão fractal do sinal. Através desse método foram analisados dois tipos de sinais sonoros: a vocalização de leitões saudáveis e de leitões acometidos por uma doença chamada Artrite Traumática; e a vocalização de suínos adultos em situações de conforto e desconforto.

Os resultados demonstram que a análise proposta atingiu bons patamares de acerto na determinação do bem-estar do animal.

**Palavras-chave:** Processamento Digital de Sinais, Suínos, Redes Neurais Artificiais, Transformada Discreta *Wavelet*



## **Abstract**

This work describes an algorithm which was created and applied to classify patterns of swine vocalizations, in order to determine the animal's welfare, since this is an issue increasingly discussed, becoming a priority in management of these animals, especially for slaughter. Thus, it is necessary to have a method in which there is no contact with the animals, avoiding modifications of the animal's behavior and, consequently, the results. The proposed approach implements the above-mentioned analysis by using an Artificial Neural Network and the Discrete Wavelet Transform. The characteristics obtained from the signals are: energies of the critical bands of the Bark scale; the differences between energies of the adjacent bands; and the fractal dimension of the signal. Through this method, two types of signals were analyzed: the vocalization of healthy piglets and sick piglets, which had Traumatic Arthritis; and the vocalization of adult pigs in situations of comfort and discomfort. The results show that the proposed method achieves good levels of accuracy in determining the animal's welfare.

**Keywords:** Digital Signal Processing, Swines, Artificial Neural Networks, Discrete Wavelet Transform

## 1. Introdução

No contexto deste trabalho, vocalização é a emissão de sons pelo animal diante de alguma situação comportamental. Neste trabalho, as situações analisadas serão a ocorrência, ou não, da doença denominada Artrite Traumática e do ambiente proporcionando conforto ou desconforto térmico aos animais.

A vocalização, segundo Appleby et al. (1999), representa uma ferramenta precisa que permite avaliar o bem-estar dos animais, uma vez que suas chamadas traduzem seu estado emocional ou suas necessidades. Manteuffel et al. (2004) concluíram que a tensão e gritos podem indicar alterações no bem-estar dos animais.

Outro ponto importante são as Redes Neurais Artificiais (RNAs) que, segundo Braga (2000), são sistemas paralelos distribuídos e compostos por unidades de processamento simples, dispostas em uma ou mais camadas e interligadas por um grande número de conexões, associadas a pesos, e que calculam determinadas funções matemáticas, tendo seu funcionamento inspirado na estrutura biológica do cérebro humano.

Também são importantes os conceitos de Transformada *Wavelet* e Dimensão Fractal, sendo que o primeiro consiste na decomposição de um sinal no tempo para o domínio da frequência; e o segundo é um valor que diz o quão uma parte específica do sinal é similar ao sinal como um todo. Desta forma, os sinais sonoros obtidos da gravação dos sons emitidos pelos suínos serão decompostos no domínio da frequência e analisados com base em uma Rede Neural Artificial que, por sua vez, possui como elementos de análise um conjunto de energias relativas ao sinal, além de sua dimensão fractal.

## 2. Objetivos

Conforme Lee et al. (2006), animais de uma mesma espécie utilizam a vocalização para a comunicação entre membros de seu grupo. Assim, como principal objetivo, este trabalho visa propor um algoritmo de análise dos sons emitidos por suíno. Analisar o som significa estudar os arquivos de áudio gerados pelas respostas dos animais diante de algum comportamento conhecido.

Testes iniciais foram realizados com sons provenientes de leitões sadios e leitões acometidos pela Artrite Traumática, com a finalidade de se identificar o padrão sonoro de cada situação para a posterior classificação de um sinal sonoro entre ambas as situações. Também foi testada a separação entre os sons advindos de uma matriz (fêmea reprodutora suína) em quatro situações de ambiente: alta e baixa entalpia e alto e baixo Índice de Temperatura de Globo e Umidade (ITGU). Neste último caso, a análise se deu na determinação das diferenças entre os níveis de ITGU e de entalpia de forma separadas. Tais índices foram escolhidos por serem índices de conforto muito utilizados em animais, principalmente em suínos, que são os animais estudados neste trabalho. Estes índices levam em consideração a temperatura do ambiente, sua umidade e radiação que, comparados com tabelas pré-estabelecidas, fornecem faixas de conforto e estresse para cada tipo de animal.

### **3. Conceitos Iniciais**

#### **3.1. Bem-estar de Suínos**

Cada vez mais, os métodos de abate de animais criados em granjas<sup>1</sup>, assim como a criação em si, estão sendo debatidos por todo o mundo. O bem-estar, mais do que nunca, está se tornando prioridade na criação dos animais. Logo, é interessante desenvolver um método para avaliar os agentes relacionados ao bem-estar do animal, como o estresse, por exemplo. Mais do que isso, o método deve possuir o menor grau possível de contato com o animal sob análise.

Nessa área, várias pesquisas têm sido realizadas com diversos tipos de animais para buscar tais métodos de avaliação. E, uma vez que a vocalização animal permite medir seu nível de estresse, é possível então analisar detalhadamente sua vocalização para obter um indicativo de seu bem-estar (Manteuffel et al., 2004).

Ambientes ou condições estressantes podem alterar o estado afetivo do animal, que, por sua vez, pode influenciar o modo como o mesmo reage aos estímulos do ambiente (Boissy et al., 2007). Logo, essas ocorrências devem ser monitoradas, pois podem representar o início de uma doença, por exemplo, o que se torna altamente preocupante quando vários animais são criados juntos, como em uma granja. Nesse sentido, e novamente segundo Manteuffel et al. (2004), os sons são

---

<sup>1</sup> Granja é uma construção fechada na qual se abrigam aves ou mamíferos (suínos especialmente) para o futuro abate.

transmitidos rapidamente e a grandes distâncias, sendo relativamente imunes a objetos que possam obstruir seu caminho. Por isso, são ideais para carregar sinais que serão enviados, por exemplo, a um alarme, que pode estar a uma grande distância.

Outro trabalho interessante foi realizado por Algers e Jensen (1985), no qual foram analisados os efeitos de um ambiente com ruído contínuo sobre leitões, sendo que três leitões foram colocados em um ambiente silencioso e outros três em um ambiente ruidoso. Para todos eles, foram reproduzidos grunhidos previamente gravados de sua matriz, os quais sofriam alterações em determinado momento. Os animais que estavam no ambiente silencioso responderam às alterações dos grunhidos reproduzidos, enquanto que os animais colocados no ambiente ruidoso não responderam à variação. Logo, foi sugerido que os animais expostos ao barulho podiam receber menos leite que os outros por terem dificuldade em responder a sinais sonoros da matriz, influenciando seu crescimento.

Na mesma linha de pesquisa sobre bem-estar animal, um trabalho sobre a adaptabilidade de suínos a diferentes temperaturas ambiente foi realizado por Hillmann et al. (2004). Alguns suínos foram testados no inverno, enquanto outros foram testados no verão. Os sons emitidos por cada grupo foram gravados à noite, evitando, assim, os efeitos e sons de atividades realizadas durante o dia. Através de redes neurais artificiais e de aplicativos como o LabVIEW®, os sons foram analisados e chegou-se à conclusão de que o ambiente afeta claramente o comportamento vocal dos animais, sendo que os mesmos emitem uma maior quantidade de sons em alta frequência quando estão desconfortáveis, ou seja, em temperaturas mais elevadas ou mais baixas. Desta forma, seria possível o monitoramento dos sons emitidos para se verificar a adaptação dos animais referente às variações de temperatura.

Ainda nesse sentido, um experimento foi realizado para se verificar a emissão de sons em altas frequências pelos suínos quando submetidos à castração a fim de se averiguar quais aspectos do processo são mais dolorosos ao animal (Taylor e Weary, 2000). Notavelmente, os animais emitiram maior quantidade de altas frequências durante a incisão, mostrando que as mesmas também podem indicar, além de desconforto, a dor sentida pelo animal.

Em um trabalho realizado por Branco et al. (2006), no qual foram analisados os choros de dor emitidos por crianças recém-nascidas, constatou-se, entre outras coisas, a ocorrência de frequências hiper-agudas, ou seja, altas frequências,

mostrando que os bebês também podem exprimir seu desconforto através da emissão de altas frequências.

Outro trabalho envolvendo altas frequências analisou os sons emitidos por leitões quando os mesmos são separados de suas matrizes (Weary et al., 1999). Descobriu-se que os leitões mais novos emitem maiores quantidades de sons em altas frequências que leitões mais velhos, pois, instintivamente, sabe-se que os leitões mais novos precisam de maiores cuidados de sua matriz, pois são mais frágeis. Outro fator que pode ter influenciado na emissão dos sons em alta frequência é a variação de temperatura sofrida pelo leitão, uma vez que sem os cuidados de sua mãe e com a redução de alimento - amamentação - o mesmo perde temperatura corporal. Desse modo, o animal fica com mais frio e desconfortável, emitindo, como supracitado, sons em alta frequência.

Em resumo, os trabalhos citados evidenciam que existe uma relação entre a análise de frequências emitidas pelos animais e o bem-estar do mesmo, podendo aquela ser utilizada para a determinação deste.

### **3.2. Digitalização de Sinais**

Sabe-se que os sons encontrados na natureza são conhecidos como contínuos, inclusive os que exprimem a comunicação entre os animais, alvo do presente trabalho. Dessa forma, existem infinitos valores entre dois valores discretos do sinal. Assim, o som que atinge o sistema auditivo dos animais causa movimentos no sistema auditivo, que, associados a mecanismos neurais, permitem distinguir dentre os diferentes tipos de sons encontrados na natureza.

Para que o som seja processado por sistemas computacionais, é preciso, inicialmente, convertê-lo para um formato de números que possa ser quantificado pelo computador, isto é, o formato digital. Tal processo é conhecido como conversão analógica-digital ou simplesmente CAD. A Figura 3.2.1 ilustra uma amostragem de um sinal sonoro hipotético.

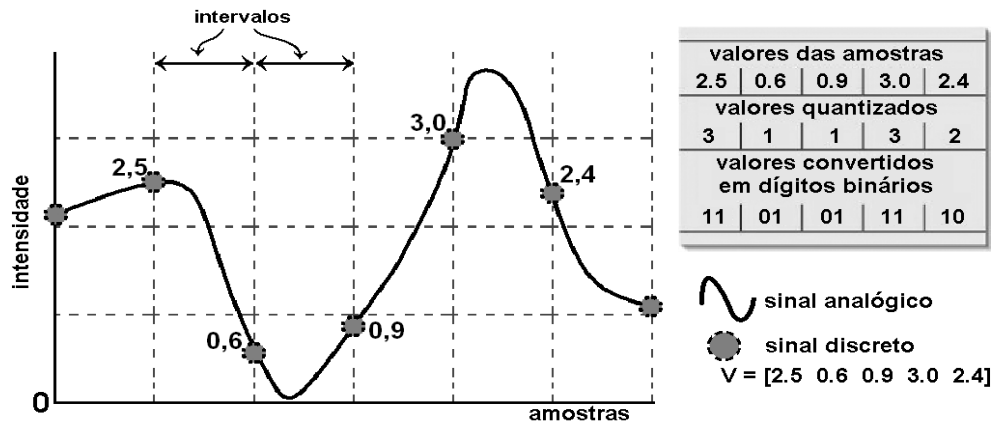


Figura 3.2.1 - Amostragem de um sinal sonoro qualquer

Como o sinal analógico possui infinitos valores nos intervalos mencionados, o processo de discretização resolve este problema atribuindo valores discretos para o novo sinal que está se formando. Desse modo, é necessária uma amostragem do sinal, que consiste em obter os valores do sinal analógico, contínuo no tempo, em períodos constantes de tempo, obtendo, assim, uma série de valores discretos.

Porém, pelo Teorema de Nyquist, a frequência de amostragem deve ser, no mínimo, duas vezes a máxima frequência encontrada no sinal, evitando assim o *aliasing* no sinal. Além disso, se o microfone utilizado possuir uma frequência de amostragem baixa, o som será capturado e convertido poucas vezes, fazendo com que o arquivo digitalizado seja ruim, como ilustra a Figura 3.2.2.

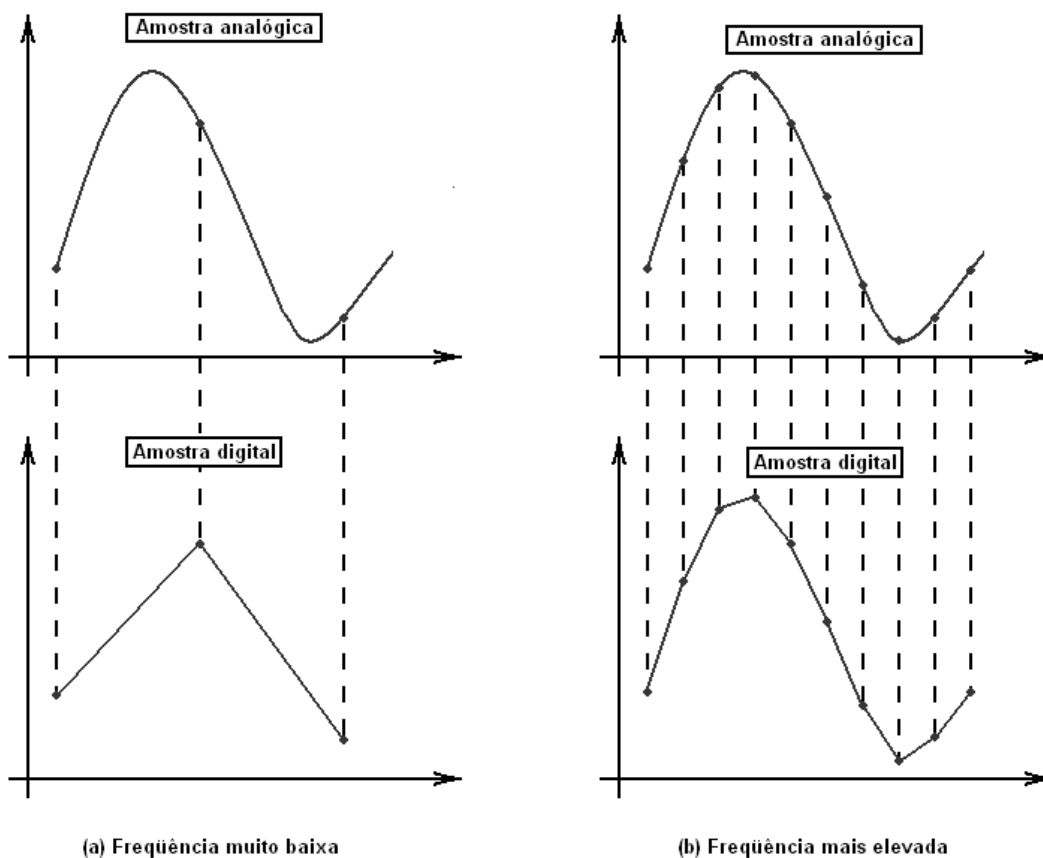


Figura 3.2.2 - Discretização do som

### 3.3. Filtros

Um conceito importante para o trabalho é a filtragem digital de sinais. Um filtro digital pode ser entendido como um sistema que, através de combinações lineares do sinal de entrada com certos coeficientes, gera uma saída com determinadas características de frequência. Alguns dos parâmetros mais relevantes, como mostrado na Figura 3.3.1, são:

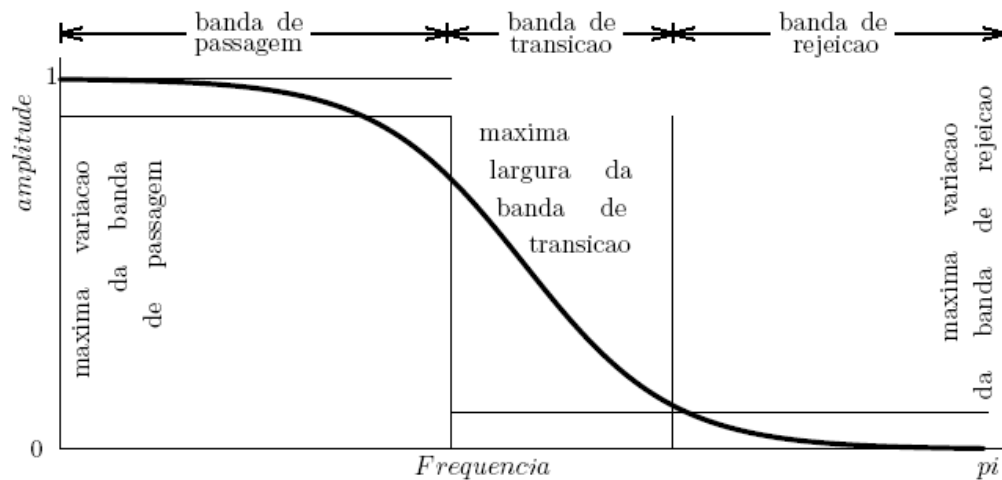


Figura 3.3.1 - Parâmetros de um filtro (Júnior, 2007)

- **Frequência de corte:** frequência na qual a saída apresenta uma atenuação de 3dB em relação à entrada e separa a banda de passagem da banda de transição.
- **Frequência de rejeição:** frequência na qual a saída apresenta um valor com considerável atenuação em relação ao do valor de entrada; normalmente 5% do valor de entrada. Analogamente, separa a banda de transição da banda de rejeição.
- **Função:** está relacionada com o tipo do filtro. Há os filtros passa-altas, passa-baixas, passa-faixas e rejeita-faixas, sendo que os filtros passa-altas permitem a passagem de altas frequências e atenua as baixas frequências. Analogamente, os filtros passa-baixas fazem o oposto: permitem a passagem de baixas frequências e atenuam as altas frequências. Do mesmo modo, os filtros passa-faixas e rejeita-faixas permitem a passagem de uma determinada faixa de frequências e atenuam todas as outras, e vice-versa, respectivamente.
- **Ordem:** é o número de pólos da função de transferência do filtro, uma vez que uma função de transferência é, basicamente, uma função que relaciona a saída de um sistema com sua entrada, descrita adiante. Quanto maior o número de pólos da função de transferência, ou seja,



quanto maior a ordem do sistema, melhor é a resposta do filtro, aproximando-se de um filtro ideal, que é aquele no qual não há banda de transição, a banda de passagem não é atenuada e a banda de rejeição é totalmente atenuada, possuindo saída nula.

Outro conceito é a determinação da energia do sinal. Tendo-se um sinal discreto, de amplitude e duração finitas, pode-se calcular sua energia,  $E(x[n])$ , como sendo

$$E = \sum_{i=0}^{N-1} x_i^2 \quad [3.3.1]$$

no qual  $x_i$  é o  $i$ -ésimo componente do sinal discreto.

Além disso, existe um importante teorema no processamento de sinais, o Teorema da Convolução, que mostra que a multiplicação de dois sinais discretos no domínio da frequência,  $H[z]$  e  $X[z]$ , corresponde à convolução dos mesmos no domínio do tempo,  $h[n]$  e  $x[n]$ , e pode ser descrita por

$$y[n] = x[n] * h[n] = \sum_{k=0}^{M-1} h_k x_{n-k} \quad [3.3.2]$$

no qual o símbolo  $*$  representa a convolução e  $M$  o número de amostras de  $x[.]$ .

### 3.4. Redes Neurais Artificiais

Conforme Silva et al. (2010), as Redes Neurais Artificiais (RNAs) podem ser entendidas como uma tentativa de simular nos computadores a estrutura e o funcionamento do cérebro, particularmente dos neurônios. Tais neurônios podem ser vistos, simplificadamente, como sendo células divididas em três partes: os dendritos, o corpo celular e o axônio, como mostrado na Figura 3.4.1, abaixo.

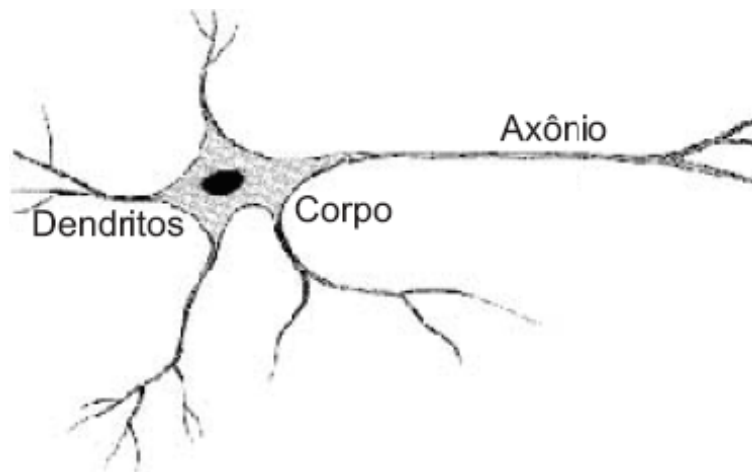


Figura 3.4.1 - Neurônio simplificado

A principal função dos dendritos é a recepção de estímulos advindos de diversos outros neurônios conectados ao mesmo. Tais estímulos são, então, processados pelo corpo celular, que produzirá um potencial de ativação para disparar um impulso elétrico pelo axônio. Por fim, seu axônio possui a função de conduzir os impulsos elétricos gerados pelo corpo celular para outros neurônios que estejam conectados ao mesmo.

Um neurônio, então, recebe pulsos como estímulo, os quais são processados e, atingido certo limiar, o neurônio realiza uma ação, que pode ser tanto a emissão de pulsos para um neurônio vizinho como a inibição do mesmo.

Estima-se que a rede neural biológica de um adulto possua cerca de 100 bilhões de neurônios, cada um fazendo conexões com outros 6.000 neurônios, perfazendo um total de 600 trilhões de sinapses (Shepherd, 1990), sendo que as sinapses correspondem às regiões de comunicação entre os neurônios. Vale ressaltar que tais conexões não são realizadas através de contato físico entre os neurônios, mas através de elementos chamados neurotransmissores, que são os responsáveis pela transmissão de impulsos elétricos entre um neurônio e outro.

O modelo de neurônio artificial mais simples foi proposto por McCulloch & Pitts e ainda é o modelo mais utilizado nas diversas redes neurais artificiais. Tal modelo pode ser visto na Figura 3.4.2.

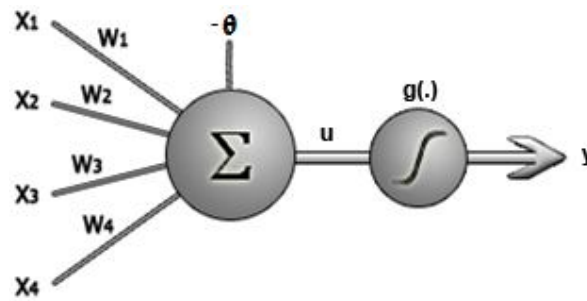


Figura 3.4.2 - Modelo de um neurônio artificial

Sua estrutura é:

- Sinais de entrada  $\{x_1, \dots, x_n\}$ : são os sinais advindos do processo a ser analisado.
- Pesos sinápticos  $\{w_1, \dots, w_n\}$ : são os valores que permitirão à rede neural avaliar a relevância de cada sinal de entrada.
- Combinador linear  $\{\Sigma\}$ : agrega todos os sinais de entrada devidamente ponderados pelos seus respectivos pesos sinápticos a fim de produzir um valor de potencial de ativação.
- Limiar de ativação  $\{\theta\}$ : valor que especifica qual o limiar que deve ser atingido pelo potencial de ativação para que a saída do neurônio seja ativada.
- Potencial de ativação  $\{u\}$ : diferença produzida entre o combinador linear e o limiar de ativação. Caso o potencial de ativação seja positivo ( $u \geq \theta$ ), o neurônio produzirá um potencial excitatório; caso contrário, o potencial será inibitório.
- Função de ativação  $\{g\}$ : sua função é modificar e limitar o valor da saída do neurônio dentro uma faixa de valores possíveis.
- Sinal de saída  $\{y\}$ : resultado do processamento do neurônio em relação às entradas fornecidas. Tal valor será propagado para os neurônios que estão seqüencialmente conectados.

Desta forma, é fácil verificar que as equações de um neurônio artificial são dadas por

$$u = \sum_{i=1}^n w_i x_i - \theta \quad [3.4.1]$$

$$y = g(u) \quad [3.4.2]$$

A função de ativação pode assumir muitas formas, sendo as mais comuns a Função Degrau, a Função Degrau Bipolar, a Função Rampa Simétrica, a Função Logística (Sigmoidal), a Função Tangente Hiperbólica e a Função Gaussiana. Esta última pode ser descrita como

$$g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}} \quad [3.4.3]$$

na qual  $c$  representa o centro da função gaussiana e  $\sigma$  denota o desvio padrão associado à mesma. A representação gráfica da função gaussiana é vista na Figura 3.4.3.

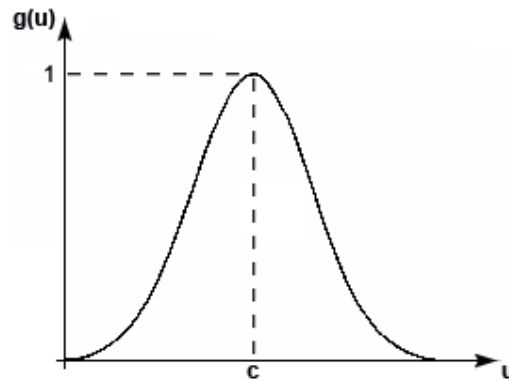


Figura 3.4.3 - Função Gaussiana

Já a rede neural artificial é o conjunto de neurônios artificiais conectados entre si. Tal arranjo pode ser dividido em três partes: a camada de entrada, a camada intermediária e a camada de saída. Os neurônios que recebem os sinais de entrada na rede constituem o que se chama de *camada de entrada*. Os neurônios que recebem como entrada as saídas daqueles da camada de entrada constituem a segunda camada e assim sucessivamente até a camada final, que é a *camada de saída*. As camadas internas que não são nem a de entrada e nem a de saída são geralmente referidas como *camadas intermediárias* ou *camadas ocultas*, conforme visto na Figura 3.4.4.

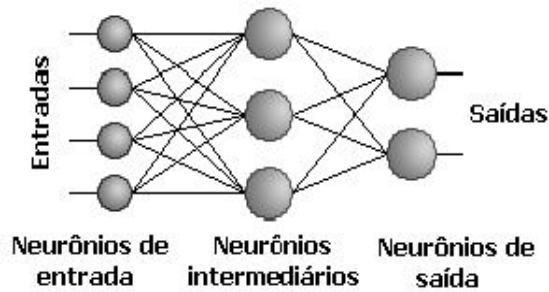


Figura 3.4.4 - Esquema de uma Rede Neural Artificial

As principais características envolvidas com a utilização de Redes Neurais Artificiais são:

- Adaptação por experiência: os parâmetros internos da rede, tipicamente seus pesos sinápticos, são ajustados através da análise de exemplos relacionados com o comportamento do processo considerado.
- Capacidade de aprendizado: uma vez que os parâmetros internos da rede são ajustados conforme novos exemplos são analisados, a rede é capaz de extrair os relacionamentos entre as variáveis existentes. Tal processo é chamado de treinamento da rede.
- Generalização: com a rede devidamente treinada, sinais ainda desconhecidos pela mesma podem ser analisados, estimando-se a solução dos mesmos através dos padrões aprendidos durante a fase de treinamento.
- Tolerância a falhas: a rede torna-se tolerante a falhas quando parte de sua estrutura é danificada. Este fato deve-se à grande quantidade de interligações entre os neurônios artificiais.
- Armazenamento distribuído: cada sinapse entre os neurônios artificiais realiza parte do processamento do sinal, de forma que o conhecimento do comportamento de determinado sinal é distribuído entre os vários neurônios artificiais que compõem a rede neural artificial.
- Facilidade de prototipagem: uma vez treinada a rede, o processamento de novos sinais se dará, geralmente, através de operações matemáticas elementares, possibilitando sua fácil prototipagem tanto em *software* como em *hardware*.

Uma importante configuração de uma RNA é aquela caracterizada pelas Funções de Base Radial (*RBF – Radial Basis Function*). Sua utilização é ampla em problemas que envolvem aproximações de funções e classificação de padrões. Por este motivo, esta foi a rede utilizada neste trabalho. Sua estrutura típica consiste na utilização de apenas uma camada oculta, como a disposta na Figura 3.4.4, e funções de ativação do tipo gaussiana. Sua arquitetura é do tipo *feedforward* de camadas múltiplas, uma vez que o sinal é transmitido unidirecionalmente da camada de entrada para a camada de saída; e porque possui mais de uma camada no total.

O treinamento de uma rede RBF é dividido em dois estágios:

### Primeiro Estágio

Ajuste dos neurônios da camada intermediária: como já mencionado, a expressão que define uma função de ativação gaussiana é dada por

$$g(u) = e^{-\frac{(u-c)^2}{2\sigma^2}} \quad [3.4.4]$$

em que  $c$  define o centro da função gaussiana e  $\sigma^2$  denota sua variância ( $\sigma$  denota o desvio padrão), a qual denota o quão o potencial de ativação ( $u$ ) está disperso em relação ao seu centro ( $c$ ).

Desta forma, considerando a equação acima, os parâmetros livres são o centro  $c$  e a variância  $\sigma^2$ . Pela configuração da rede RBF, o centro  $c$  está diretamente associado aos seus próprios pesos, enquanto os valores de entrada estão associados com o próprio sinal da entrada da rede. Desta forma, a saída de cada neurônio  $j$  da camada intermediária é expressa por

$$g_j^{(1)}(u_j^{(1)}) = g_j^{(1)}(x) = e^{-\frac{\sum_{i=1}^n (x_i - w_{ji}^{(1)})^2}{2\sigma_j^2}} \quad [3.4.5]$$

no qual  $x$  representa o próprio vetor de entrada; o índice  $^{(1)}$  é relativo à camada intermediária;  $j = 1, \dots, n_1$  (sendo  $n_1$  o número de neurônios na camada intermediária).

Pode-se perceber, através da equação 3.4.4, que quanto mais próximo uma determinada amostra está do centro da gaussiana, maior será sua contribuição para o valor do potencial de ativação, sendo este valor o mesmo para qualquer amostra situada na mesma distancia radial ao centro  $c$ .

Tratando-se de reconhecimento de padrões, as fronteiras delimitadoras – chamadas fronteiras de separabilidade: fronteira entre os padrões analisados – são definidas como campos receptivos hiperbólicos, conforme mostrado pelas Figura 3.4.5 e Figura 3.4.6 para um problema constituído de duas entradas  $x_1$  e  $x_2$ .

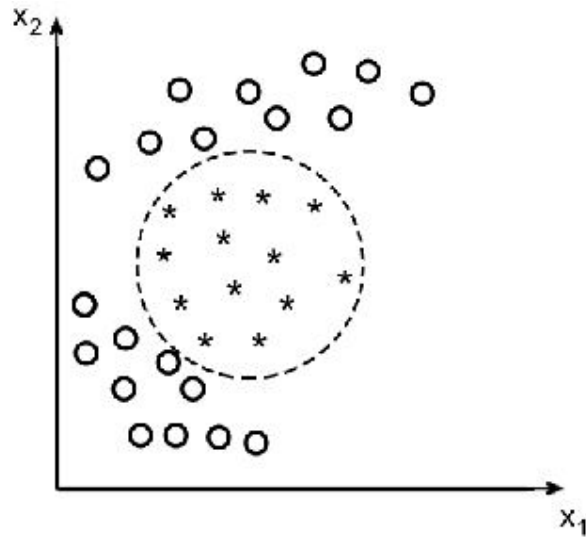


Figura 3.4.5 - Função de base radial do tipo gaussiana

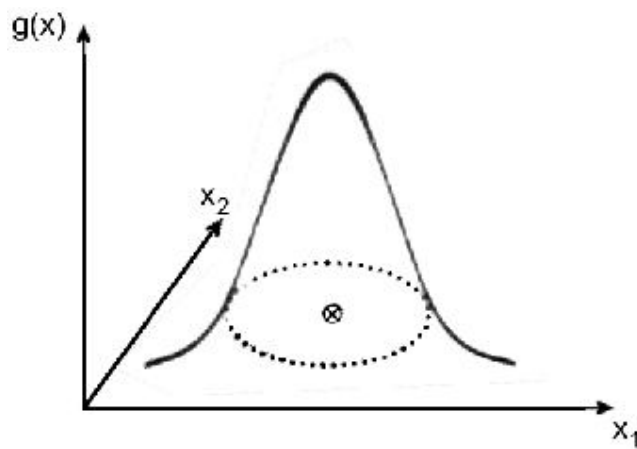


Figura 3.4.6 - Fronteira de separabilidade RBF

Dessa forma, a principal função da camada oculta é posicionar o centro de suas gaussianas da forma mais apropriada possível. O pseudocódigo para este estágio de treinamento de uma rede RBF é tal como se segue:

### **Início {Algoritmo RBF – Primeiro Estágio de Treinamento}**

- <1> Obter o conjunto de treinamento  $\{x^{(k)}\}$ ;
- <2> Iniciar o vetor de pesos de cada neurônio da camada intermediária com os valores das  $n_1$  primeiras amostras de treinamento;
- <3> Repetir as instruções:
  - <3.1> Para todas as amostras de treinamento  $\{x^{(k)}\}$ , fazer:
    - <3.1.1> Calcular as distâncias euclidianas entre  $x^{(k)}$  e  $w_{ji}^{(1)}$ , considerando-se cada  $j$ -ésimo neurônio por vez;
    - <3.1.2> Selecionar o neurônio  $j$  que contenha a menor distância com o intuito de agrupar a referida amostra junto ao centro mais próximo;
    - <3.1.3> Atribuir a amostra  $x^{(k)}$  ao grupo  $\Omega^{(j)}$ ;
  - <3.2> Para todos  $w_{ji}^{(1)}$ , onde  $j=1, \dots, n_1$ , fazer:
    - <3.2.1> Ajustar  $w_{ji}^{(1)}$  de acordo com as amostras em  $\Omega^{(j)}$ :
$$w_{ji}^{(1)} = \frac{1}{m^{(j)}} \sum_{x^{(k)} \in \Omega^{(j)}} x^{(k)}$$

$\{m^{(j)}\}$  é o número de amostras em  $\Omega^{(j)}\}$

Até que: não haja mudanças nos grupos  $\Omega^{(j)}$  entre as iterações;
- <4> Para todos  $w_{ji}^{(1)}$ , onde  $j = 1, \dots, n_1$ , fazer:
  - <4.1> Calcular a variância de cada uma das funções de ativação gaussianas pelo critério da distancia quadrática média:
$$\sigma^2 = \frac{1}{m^{(j)}} \sum_{x^{(k)} \in \Omega^{(j)}} \sum_{i=1}^n (x_i^{(k)} - w_{ji}^{(1)})^2$$

### **Fim {Algoritmo RBF – Primeiro Estágio de Treinamento}**

### **Segundo Estágio**

Ajuste dos neurônios da camada de saída: após a etapa anterior ser concluída, o segundo estágio de treinamento deve ser executado. Neste estágio, utiliza-se a função linear como função de ativação para os neurônios de saída, de modo que



$$y_i = g_j^{(2)}(u_j^{(2)}) = u_j^{(2)}, \quad [3.4.6]$$

no qual  $j=1, \dots, n_2$  e o índice  $(2)$  está relacionado com a camada de saída.

Desta forma, a camada de saída realiza apenas uma combinação linear das funções de ativação gaussianas produzidas na camada intermediária.

O pseudocódigo deste estágio é como se segue.

#### **Início {Algoritmo RBF – Segundo Estágio de Treinamento}**

- <1> Obter o conjunto original de amostras de treinamento  $\{x^{(k)}\}$ ;
  - <2> Obter o vetor de saída desejada  $\{d^{(k)}\}$  para cada amostra;
  - <3> Iniciar  $w_{ji}^{(2)}$  com valores aleatórios pequenos;
  - <4> Especificar a taxa de aprendizagem  $\{\eta\}$  e precisão requerida  $\{\varepsilon\}$ ;
  - <5> Para todas as amostras  $\{x^{(k)}\}$ , fazer:
    - <3.1> Obter os valores  $g_j^{(1)}$  em relação a  $x^{(k)}$ ; {conforme explicado anteriormente}
    - <3.2> Assumir  $z^{(k)} = [g_1^{(1)} \ g_2^{(1)} \dots \ g_{n_1}^{(1)}]^T$ ; {pseudoamostras}
  - <6> Inicializar o contador de número de épocas; {época  $\leftarrow 0$ }
  - <7> Repetir as instruções:
    - <7.1>  $E_M^{anterior} \leftarrow E_M$ ;
    - <7.2> Para todos os pares de treinamento  $\{z^{(k)}, d^{(k)}\}$ , fazer:
      - Ajustar  $w_{ji}^{(2)}$  e  $\theta_j$  através do algoritmo *backpropagation*;
    - <7.3>  $E_M^{atual} \leftarrow E_M$ ;
    - <7.4> época  $\leftarrow$  época + 1;
- Até que  $|E_M^{atual} - E_M^{anterior}| \leq \varepsilon$

#### **Fim {Algoritmo RBF – Segundo Estágio de Treinamento}**

A variável *época* pode ser utilizada como critério de parada para problemas em que a precisão especificada não pode ser alcançada;  $E_M$  representa o erro quadrático médio, que é definido por:

$$E_M = \frac{1}{p} \sum_{k=1}^p E(k) \quad [3.4.7]$$

No qual  $p$  é o número de amostras do conjunto de treinamento e  $E(k)$  representa o erro quadrático obtido por:

$$E(k) = \frac{1}{2} \sum_{j=1}^{n_2} \left( d_j(k) - y_j^{(2)}(k) \right)^2 \quad [3.4.8]$$

no qual  $y_j^{(2)}(k)$  é o valor produzido pelo  $j$ -ésimo neurônio de saída da rede considerando-se a  $k$ -ésima amostra de treinamento, enquanto  $d_j(k)$  é o seu respectivo valor desejado. Já o algoritmo *backpropagation* é um método de se calcular os pesos sinápticos e os limiares de ativação de forma rápida através de gradientes, porém seu funcionamento está além do escopo deste trabalho<sup>2</sup>.

Uma vez que estas etapas estiverem concluídas, o treinamento está completo. Dessa forma, a rede neural está pronta para operar, ou seja, classificar novas amostras que são apresentadas à sua entrada. Os passos dessa operação estão mostrados abaixo.

#### **Início {Algoritmo RBF – Fase de Operação}**

- <1> Apresentar uma amostra  $\{x\}$ ;
- <2> Assumir os parâmetros  $w_{ji}^{(1)}$ ,  $\sigma$ ,  $w_{ji}^{(2)}$  e  $\theta_j$  já ajustados durante os estágios de treinamento;
- <3> Executar as seguintes instruções:
  - <3.1> Obter  $g_j^{(1)}$ ;
  - <3.2> Obter  $u_j^{(2)}$ ;
  - <3.3> Obter  $y_j$ ;
- <4> Disponibilizar as saídas da rede mediante os valores contidos em  $y_j$ ;

#### **Fim {Algoritmo RBF – Fase de Operação}**

Segundo Jain et al. (1996), as RNAs são a base de inúmeros avanços no desenvolvimento de sistemas inteligentes. Tais redes são bem adequadas para tarefas

---

<sup>2</sup> Para maiores detalhes do funcionamento desse método, vide Silva et al. (2010)

que incluam problemas como reconhecimento de padrões, predição, otimização, memória associativa e controle. Ou seja, são extremamente úteis a este trabalho, uma vez que se busca reconhecer os padrões entre os diferentes estados de um animal como, por exemplo, se está estressado ou não. O uso de RNAs é vasto, podendo ser utilizadas em qualquer situação que necessite de algumas das características supracitadas, como o reconhecimento de padrões. Por isso, são usadas desde a compressão de áudio até em diagnósticos médicos, assim como em processamento de sinais.

Porém, segundo Smith (1998), o reconhecimento automatizado da fala humana é um exemplo clássico de uma ação que o cérebro humano realiza facilmente, mas que os computadores possuem enormes problemas para fazê-lo. Os computadores digitais podem guardar e acessar vastas quantidades de dados, executar cálculos matemáticos em enorme velocidade e fazer tarefas repetitivas sem se aborrecer ou tornar-se ineficiente. Porém, são altamente ineficientes quando afrontados por dados sensoriais sem tratamento. O Processamento Digital do Sinal geralmente aproxima este problema em dois passos: a extração das características seguida pelo reconhecimento dos padrões. Os pesos requeridos para a Rede Neural reconhecer um padrão são encontrados através de um algoritmo de aprendizagem, junto com exemplos de como o sistema deveria operar.

Tem-se aí um problema, que é a dificuldade de reconhecer os padrões na fala humana, e também na “fala animal”. Para contornar a situação, a extração das características é feita através da Transformada Discreta *Wavelet*, descrita abaixo. Mas, para um melhor resultado, deve-se entregar à RNA uma quantidade significativa de exemplos para cobrir a maior gama de possibilidades possível.

### **3.5. Transformada *Wavelet***

No início do século 19, Jean-Baptiste Joseph Fourier mostrou que qualquer função periódica pode ser escrita como uma soma de funções periódicas seno e cosseno. Em outras palavras, um sinal periódico pode ser decomposto em frequências específicas que constituem o mesmo. Esta é a definição informal da série de Fourier, que, formalmente, pode ser representada da seguinte forma:

Sendo

$$f(t + 2L) = f(t), \quad c \leq t \leq c + 2L \quad [3.5.1]$$

Então

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[ a_n \cdot \cos\left(\frac{n\pi t}{L}\right) + b_n \cdot \text{sen}\left(\frac{n\pi t}{L}\right) \right] \quad [3.5.2]$$

Em que  $a_0$ ,  $a_n$  e  $b_n$  são dados por

$$a_0 = \frac{1}{L} \int_c^{c+2L} f(t) dt \quad [3.5.3]$$

$$a_n = \frac{1}{L} \int_c^{c+2L} f(t) \cos\left(\frac{n\pi t}{L}\right) dt \quad [3.5.4]$$

$$b_n = \frac{1}{L} \int_c^{c+2L} f(t) \text{sen}\left(\frac{n\pi t}{L}\right) dt \quad [3.5.5]$$

Porém, há a limitação das Séries de Fourier serem aplicadas apenas a funções periódicas. Para sinais não periódicos pode-se recorrer à Transformada de Fourier, que consiste em uma transformada integral que expressa um dado sinal em termos de funções senoidais, assim como suas respectivas amplitudes. Da mesma forma, sua definição é a seguinte:

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad [3.5.6]$$

$$f(t) = F^{-1}(F(\omega)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{it\omega} d\omega \quad [3.5.7]$$

Desta forma, uma dada função pode ser representada através de seus valores no tempo – domínio do tempo – ou através de suas frequências constituintes – domínio da frequência. Tais domínios são mais ou menos vantajosos para determinados tipos de análises, sendo muito comum a passagem de um determinado sinal de um domínio para o outro através de séries ou Transformadas de Fourier.

Todavia, como pode ser percebido, a representação do sinal em determinado domínio não fornece os resultados encontrados no outro domínio. Desta forma, quando um sinal é analisado através da Transformada de Fourier, ou são conhecidos seus valores temporais ou suas frequências constituintes, de modo que não é possível determinar quando (em que intervalo de tempo) tais frequências aparecem no sinal,

uma vez que a relação entre as frequências presentes no sinal e o domínio temporal é inexistente (Fournier, 1995; Gomes et al., 1987; Gasquet et al., 2000).

Para tentar evitar este problema foi proposta a Transformada de Fourier de Tempo Curto (*STFT – Short Time Fourier Transform*), também conhecida como Transformada de Gabor (Gabor, 1953). A idéia principal desta transformada é a introdução de um parâmetro de frequência local, ou seja, o sinal é analisado por partes, dentro de “janelas”, sendo aplicada a transformada de Fourier somente no sinal contido na janela analisada. Formalmente, a transformada local observa  $f(t)$  durante uma janela  $W(t)$  centrada no instante  $t$  e de extensão limitada, ou seja,

$$STFT(\omega, \tau) := \int_{-\infty}^{\infty} f(t)W(t - \tau)e^{-j\omega t} dt \quad [3.5.8]$$

Contudo, uma vez fixada a janela  $W(t)$ , a resolução permanece constante em todo o plano tempo-frequência. Todavia, considerando por simplificação apenas as ondas fundamentais (senóides) de um sinal, pode-se perceber, intuitivamente, que dois tons de 1 kHz e 2 kHz são bem mais distintos que outro par de tons de 1,001 MHz e 1,002 MHz, embora a distância em ambos os casos seja de 1 kHz. Conforme a frequência analisada é incrementada, a distinção entre dois sinais próximos em 1 kHz torna-se mais árduo. Distinguir 1 GHz de 1,000001 GHz é uma tarefa relativamente difícil. Logo, trata-se de um problema de valor relativo: quanto 1 kHz significa no sinal analisado?

Para resolver este inconveniente, é utilizada a Análise Multirresolucional (AMR) que, como o nome diz, é a análise na qual o sinal é analisado com diferentes resoluções para diferentes frequências (Gomes et al., 1987). Uma comparação de três frequências na resolução no plano tempo-frequência pela Análise de Gabor (STFT) e pela Análise Multirresolucional pode ser vista na Figura 3.5.1, abaixo.

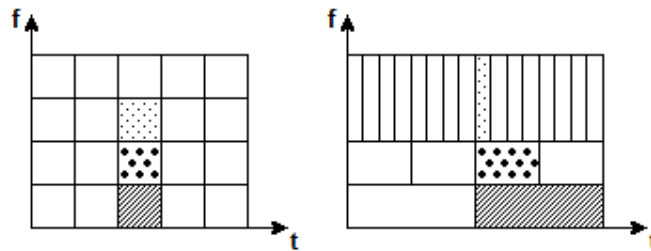


Figura 3.5.1 - Comparação entre as análises (a) de Gabor (STFT) e (b) Multirresolucional (Oliveira, 2007)

Da mesma forma, uma Transformada *Wavelet* (*WT – Wavelet Transform*) é a aplicação de uma determinada janela ao sinal, assim como a STFT. A diferença entre ambas as transformadas é que a *Wavelet* utiliza a análise multirresolucional, de modo que a janela é expandida e comprimida de acordo com a frequência em análise, ou seja, possui um parâmetro de escala ( $a$ ), além do parâmetro de descolamento da janela ( $\tau$ ). Formalmente,

$$CWT(a, \tau) := \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t) \psi^* \left( \frac{t - \tau}{a} \right) dt \quad [3.5.9]$$

Na qual CWT denota a Transformada *Wavelet* Contínua (*Continuous Wavelet Transform*) e  $\psi(t)$  denota a janela utilizada, chamada de *wavelet* básica, uma vez que todas as outras janelas são versões escalonadas (expandidas ou comprimidas) e transladadas desta *wavelet* (Morettin, 1997).

Por essa definição, se os parâmetros  $a$  e  $\tau$  forem discretizados, as janelas utilizadas não serão escalonadas e transladadas continuamente, mas em intervalos discretos, dando origem à Transformada Discreta *Wavelet* (*DWT – Discrete Wavelet Transform*).

Sendo assim, as *wavelets* podem ser interpretadas como as transformadas lineares locais geradas por um banco de filtros passa-faixas de fator de qualidade constante (Meyer et al., 1987), ou seja, uma série de filtros cuja banda passante é proporcional à frequência central do mesmo ( $Q = \Delta f / f = cte$ ).

Em oposição à Transformada *Wavelet*, a STFT, por utilizar janelas fixas para todas as frequências analisadas, possui banda passante constante para qualquer frequência do sinal. A comparação da análise espectral entre essas duas transformadas pode ser vista na Figura 3.5.2.

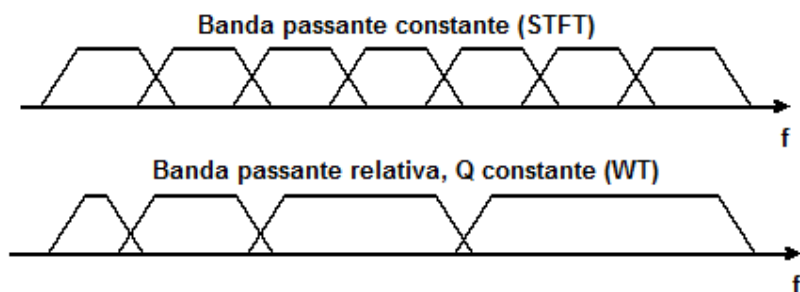


Figura 3.5.2 - Comparação da análise espectral entre (a) STFT e (b) WT (Oliveira, 2007)

Essa análise através de um banco de filtros é conhecida como Codificação em Sub-bandas (*Subband Coding*) e os filtros utilizados no cálculo da DTW são filtros passa-baixas e filtros passa-altas calculados de forma a dividir o espectro do sinal ao meio, ou seja, suas frequências de corte são ajustadas na metade da máxima frequência presente no sinal. Com isso, as componentes resultantes da filtragem passa-altas contêm as frequências mais altas do sinal, que já fornecem as características de forma detalhada, conforme a AMR. Já, as componentes resultantes da filtragem passa-baixas contêm as baixas frequências do sinal, de modo que precisam ser analisadas novamente para se obter informações mais acuradas destas frequências (Mallat, 2000). Para isso, basta passar essas componentes de baixa frequência por um novo par de filtros passa-altas e passa-baixas para novamente dividir este espectro em duas partes, e assim sucessivamente. Cada vez que o processo de filtragem é executado diz-se que se tem um novo nível de decomposição.

Como é possível perceber, cada sinal gerado pelo DWT possui apenas metade das frequências presentes no sinal anterior. Tal processo é conhecido como sub-amostragem por 2 (*down-sampling*). Os coeficientes resultantes dos filtros passa-baixas e passa-altas são conhecidos por coeficientes de aproximação e detalhamento, respectivamente.

Um sinal contendo  $n$  amostras gera uma transformada *wavelet* de igual número de amostras, sendo composta por uma sequência de coeficientes. Esta sequência é iniciada pelos coeficientes de aproximação do último nível de decomposição, seguidos pelos coeficientes de detalhamento dos níveis intermediários e finalizando com os coeficientes de detalhamento do primeiro nível de decomposição, como mostrado na Figura 3.5.3 (Junior, 2007).

Para se realizar esse processo recursivo até o último nível possível, é necessário, obviamente, que o número de amostras  $n$  do sinal discreto original esteja na forma  $2^m$ , ou seja, uma potência de dois, pois em cada decomposição o sinal é dividido em duas partes. Logo, a quantidade de decomposições possíveis de um sinal com  $n$  amostras é

$$2^m = n \Rightarrow m = \frac{\log(n)}{\log(2)} \quad [3.5.10]$$

Assim, um sinal contendo  $n$  amostras pode ser decomposto em, no máximo,  $m = \frac{\log(n)}{\log(2)}$  níveis.

Desse modo, pode-se entender o processo da DWT em cada nível como sendo um par de filtros, sendo

$$y_{\text{passa-baixas}}[n] = x[n] * h[n] = \sum_{k=0}^{n-1} h_k x_{2n-k} \quad [3.5.11]$$

$$y_{\text{passa-altas}}[n] = x[n] * g[n] = \sum_{k=0}^{n-1} g_k x_{2n-k} \quad [3.5.12]$$

onde, obviamente,  $h[n]$  e  $g[n]$  são os filtros passa-baixas e passa-altas, respectivamente.





### 3.6. Bandas Críticas

O ouvido humano é um sistema que detecta as várias frequências presentes em um sinal sonoro. Sua faixa de percepção é dada como entre 20 Hz e 20 kHz, porém, conforme a pessoa envelhece, esta faixa diminui, geralmente degradando primeiramente os sons com maior frequência.

Além disso, o ouvido humano não é perfeito, e como tal apresenta algumas características que são estudadas através da Psicoacústica, que estuda como as pessoas interpretam os sons, relacionando as características físicas destes com o modo em que são interpretadas pela pessoa.

Na porção interna do ouvido humano existe uma estrutura na forma de um tubo ósseo enrolado em espiral chama cóclea, a qual é preenchida por um fluido e possui um grande número de células nervosas em forma de cabelo em uma membrana, chamada membrana basilar. Estas células são de tamanhos e elasticidade variados, que entram em movimento de forma diferente para determinadas frequências do sinal sonoro que atravessam o fluido da cóclea. Tal movimento é relativo à frequência de vibração específica de cada célula. Essa vibração causa um impulso elétrico correspondente à amplitude daquela frequência específica.

Porém, o ouvido não consegue distinguir muito bem entre frequências próximas, de forma que, através de experimentos, foram definidas bandas críticas, as quais consistem em faixas de frequências com uma particularidade: se duas frequências estiverem na mesma banda crítica, somente a de maior intensidade será interpretada (Barbedo, 2004).

Dessa forma, e como a sensibilidade do ouvido é dependente da frequência, a largura das bandas é variável de acordo com as frequências em questão. Tais bandas foram mapeadas através de experimentos com os quais foi criada a Escala de Bark, proposta por Eberhard Zwicker em 1961 (Zwicker, 1961). Tal escala pode ser vista na Tabela 3.6.1.

No presente trabalho, as bandas críticas foram usadas como uma tentativa de fazer com que o sistema proposto possa realizar a avaliação de determinado som animal de forma mais similar a análise feita pelo ouvido humano.

Bark	Frequência			Banda
	Mínima [Hz]	Máxima [Hz]	Central [Hz]	
0	20	100	60	80
1	100	200	150	100
2	200	300	250	100
3	300	400	350	100
4	400	510	455	110
5	510	630	570	120
6	630	770	700	140
7	770	920	845	150
8	920	1080	1000	160
9	1080	1270	1175	190
10	1270	1480	1375	210
11	1480	1720	1600	240
12	1720	2000	1860	280
13	2000	2320	2160	320
14	2320	2700	2510	380
15	2700	3150	2925	450
16	3150	3700	3425	550
17	3700	4400	4050	700
18	4400	5300	4850	900
19	5300	6400	5850	1100
20	6400	7700	7050	1300
21	7700	9500	8600	1800
22	9500	12000	10750	2500
23	12000	15500	13750	3500
24	15500	20000	17750	4500

Tabela 3.6.1 – Escala Bark

### 3.7. Dimensão Fractal

Harold E. Hurst foi um hidrólogo designado para trabalhar em uma represa no rio Nilo. Seu trabalho era estudar uma política de fluxos a ser adotada na represa para que esta não ficasse vazia e nem transbordasse. Para isso, suposições sobre o fluxo pluvial eram necessárias e, na época, era comum a suposição de que as chuvas seguiam o acaso.

Nesse ambiente, Hurst desenvolveu uma nova técnica seguindo do pressuposto que o fluxo pluvial seguia um padrão não aleatório. Desta forma, foi criado o Expoente de Hurst,  $H$ , que serve para distinguir uma série aleatória de uma

não aleatória, ou seja, tal expoente fornece informações sobre a correlação e persistência em uma série temporal. Com isto, Hurst descobriu que as chuvas seguiam uma tendência com ruído, que poderiam ser medidos de forma estatística (SOTERRONI et al., 2008).

Este expoente é dado no intervalo entre 0 e 1, significando:

- $0 \leq H < 0,5$ : o sinal é dito anti-persistente, ou seja, a probabilidade do sinal se alternar (um valor “negativo” ser seguido de um valor “positivo”) é maior que cinquenta por cento.
- $H = 0,5$ : o sinal é completamente aleatório.
- $0,5 < H \leq 1$ : o sinal é dito persistente, existindo uma probabilidade maior que cinquenta por cento do sinal se repetir.

Dessa forma, por fornecer informações sobre correlação de uma série, o expoente de Hurst é relacionado com a dimensão fractal, a qual é um valor que indica o quão uma parte de um dado sinal é similar ao sinal como um todo.

Por fim, o expoente de Hurst pode ser calculado de diversas formas, incluindo através de *Transformadas Wavelets*. Através destas, o expoente pode ser calculado através do algoritmo denominado *Power Spectrum*:

#### **Início {Algoritmo Hurst – Cálculo do Expoente Hurst}**

- <1> Aplicar no sinal a Transformada Discreta *Wavelet* no maior nível de decomposição possível;
- <2> Elevar ao quadrado cada coeficiente obtido;
- <3> Normalizar o resultado obtido através do maior valor encontrado;
- <4> Aplicar o logaritmo na base 2 em cada valor obtido no passo anterior;
- <5> Calcular o somatório do número de elementos que compõem o sinal ( $sx$ ) e do quadrado dos mesmos ( $sx^2$ ), ou seja,

$$sx = \sum_{i=1}^n i \quad sx^2 = \sum_{i=1}^n i^2$$

onde  $n$  é o número de elementos do sinal;

- <6> Calcular o somatório do sinal encontrado no passo <4>  $\{x\}$  e o somatório do mesmo multiplicado pelo seu índice, ou seja,

$$sy = \sum_{i=1}^n x_i \quad sxy = \sum_{i=1}^n (i * x_i)$$

no qual  $x_i$  corresponde ao  $i$ -ésimo valor de  $x$ , o qual corresponde ao resultado encontrado no passo <4>;

<7> Calcular o Expoente de Hurst através da seguinte equação:

$$H = \left| \frac{\frac{n * \sum xy - \sum y * \sum x}{n * \sum x^2 - (\sum x)^2} - 1}{2} \right|$$

**Fim {Algoritmo Hurst – Cálculo do Expoente Hurst}**

#### **4. Descrição das Atividades**

A pesquisa envolveu duas etapas, sendo que a primeira foi adquirir arquivos de áudio. Esse trabalho foi realizado por uma médica veterinária pesquisadora do NUPEA (Núcleo de Pesquisa em Ambiência, ESALQ, USP). O experimento foi conduzido em uma granja comercial de produção de suínos (Granja Querência), localizada no município de Elias Fausto, no interior do Estado de São Paulo. Essa granja se encontra a 23°12'10" de latitude sul e a 47°17'35" de longitude oeste e altitude média de 521 m, no centro de uma região chamada Zona de Depressão Periférica do Estado de São Paulo entre o Planalto Atlântico e a escarpa da Serra de Botucatu, com o clima caracterizado como mesotérmico.

Para a realização deste experimento, foram utilizados 12 leitões F1 (Landrace X Large White) em fase de amamentação, de ambos os sexos, com idades variando entre 10 e 15 dias de vida. Os animais foram escolhidos a partir de avaliações clínicas elaboradas por um médico veterinário, permitindo assim selecionar animais doentes e animais sadios. Sendo assim, foram selecionados seis animais doentes e seis animais sadios, que eram mantidos em celas parideiras junto com suas respectivas matrizes e demais animais da ninhada, tendo acesso livre a abrigos escamoteadores.

Os animais doentes apresentavam sinais clínicos semelhantes e indicativos da ocorrência de artrite traumática, ou seja, dificuldade em se locomover, claudicação, aumento de volume e temperatura em uma ou mais articulações dos membros, aumento sistêmico de temperatura, prostração e lesões traumáticas na porção de pele que recobre a articulação, conforme visto na Figura 4.1. Os animais classificados como sadios não apresentaram nenhum destes sinais clínicos.



Figura 4.1 - Articulação de leitões acometidos por artrite

O som emitido pelos animais foi coletado por um gravador digital da marca Panasonic® RR-US395, que pode ser visto na Figura 4.2, posicionado a uma distância de aproximadamente 15 cm da boca dos animais. Para isto, os animais foram segurados pelos veterinários, que gravaram as vocalizações. Os sons foram gravados com uma taxa de amostragem de 22050 Hz e resolução de 16 bits.



Figura 4.2 - Gravador Digital Diretivo Panasonic®RR-US395

Os leitões que apresentaram artrite foram segurados pelo corpo e os sons emitidos coletados pelo gravador. O toque sobre o corpo do animal funcionou como um estímulo para que ele vocalizasse. Após essa coleta, o animal foi marcado com uma tinta não-tóxica para posterior identificação e foi recolocado junto aos demais animais. Os seis leitões sem sinais clínicos de artrite selecionados para o experimento, portanto classificados como animais saudáveis, também foram submetidos ao mesmo procedimento para recuperar as chamadas emitidas quando segurados. A cada momento em que os animais eram segurados, uma faixa de gritos sequenciais era registrada contendo diversos gritos de suínos.

Através do software livre Audacity 1.2.6, todos os arquivos foram carregados e ouvidos separadamente. O objetivo foi montar um resumo das condições de cada

arquivo. Isso se deve porque nem todos os áudios possuíam boas condições de análise, sendo que muitos deles carregaram consigo ruídos que poderiam comprometer a análise do arquivo, como gritos excessivos dos demais filhotes, voz humana em concorrência com os grunhidos dos animais, dentre outros. Após todos os arquivos terem sido ouvidos, foram selecionados vinte e dois deles para posterior processamento, trabalho este realizado utilizando-se do software Audacity, como visto na Figura 4.3. Tais arquivos foram selecionados por apresentarem o menor ruído ambiente.



Figura 4.3 - Separação de cada grito

Foi então criado um algoritmo para, a partir das amostras brutas relativas aos sinais de áudio capturados, criar um vetor de energias referente às 23 primeiras bandas críticas da escala Bark, já percorrido acima. Foram utilizadas apenas 23 bandas por causa da máxima frequência do sinal, 11025 Hz, uma vez que a frequência de amostragem do gravador é 22050 Hz e, pelo Teorema de Nyquist, a frequência de amostragem deve ser pelo menos o dobro da maior frequência presente no sinal.

Para isso, o algoritmo calcula a Transformada Discreta *Wavelet* do sinal de entrada, extraindo dela os seguintes parâmetros:

- i. 23 valores de energia, cada um correspondente a uma banda crítica;
- ii. 22 valores das derivadas discretas das energias anteriores;
- iii. a dimensão fractal do sinal.

Com isto, cada sinal analisado gera um vetor de 46 valores (23+22+1), independentemente da duração do grito do animal, que são salvos para serem

utilizados como a entrada de uma Rede Neural Artificial, que é treinada para reconhecer os padrões.

Cada arquivo selecionado, contendo um grito cada, passou pelo *software* acima mencionado, sendo processados todos os arquivos sonoros criados a partir das vinte e duas gravações separadas. Com isto, foram analisados um total de 237 gritos de animais sadios e 248 de animais doentes, perfazendo a mesma quantia de arquivos criados, os quais possuem o vetor de características supracitado. Tais gritos foram separados dos vinte e dois arquivos selecionados, uma vez que cada um desses arquivos possuía vários gritos do animal.

Uma vez caracterizado cada grito, foram separados, para realizar o teste, vinte e um arquivos contendo os vetores relativos aos animais sadios e oito de animais doentes, sendo processados mais dez sons aleatórios obtidos de um arquivo sonoro não utilizado, contendo gritos de um animal doente. Os demais foram submetidos a uma Rede Neural Artificial, a fim de treiná-la para que esta possa reconhecer os padrões referentes aos dois tipos de entrada: o animal sadio e o animal doente. A Rede Neural Artificial foi programada, assim como a Transformada Discreta *Wavelet*, em linguagem C++.

Os arquivos contendo os vetores foram com o padrão Sadio\_X e Doente\_X, relativos aos animais sadios e doentes, respectivamente, onde “X” representa o número do arquivo. O arquivos de teste foram nomeados Teste\_X, seguindo o mesmo padrão.

Uma vez treinada, foram passados pela RNA os arquivos anteriormente separados para teste e que, por isso, não entraram no treinamento da mesma. O teste possui a finalidade de verificar a eficiência do treinamento e reconhecimento dos padrões.

Na fase de operação, a RNA terá como resultado exatamente o valor associado ao padrão, quando tiver como entrada um dos padrões a qual foi treinada pra reconhecer. Quando o valor de entrada for diferente dos usados em seu treinamento, a rede retornará um valor tão mais próximo de um dos valores associado a um padrão quanto mais similar a entrada for deste padrão.

Dessa forma, pode-se dizer que uma entrada qualquer está próxima ou não de quaisquer padrões pré-estabelecidos no treinamento da rede, que é o intuito deste trabalho.



Em suma, de forma esquemática, a metodologia das etapas envolvidas está representada na Figura 4.4.

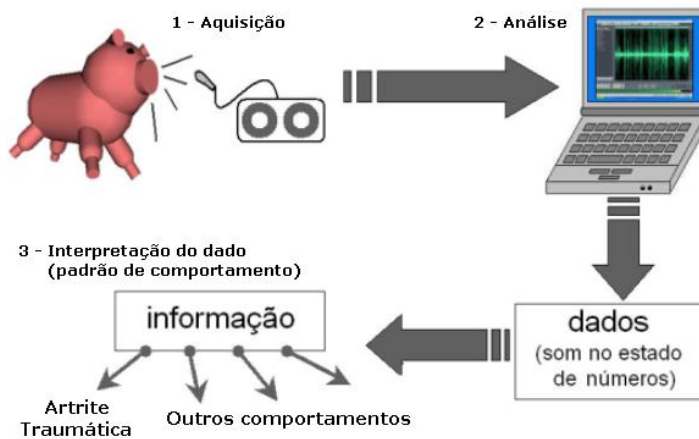


Figura 4.4 - Metodologia de aquisição e análise dos sinais de áudio

Já para os sons das matrizes, foram processados, utilizando os mesmos programas utilizados nos sons anteriores, 24 arquivos, sendo 12 de matrizes amamentando seus filhotes em um ambiente de conforto térmico e 12 de matrizes amamentando em um ambiente de desconforto térmico. Por serem poucos sons, foram separados apenas dois arquivos de cada situação para testar a rede, uma vez que, deste modo, ainda restam dez amostras de cada situação para treinar a rede. Cada arquivo de áudio gerou um arquivo contendo os vetores das características supracitadas, exatamente como foi realizado com os sons anteriormente descritos. O procedimento dessa etapa é análogo ao procedimento anterior.

## 5. Resultados

Processar uma amostra de som é carregar o referido arquivo de áudio e, na sequência, proceder com lógica matemática suficiente para alterar o comportamento do sinal primitivo.

Dessa forma, os resultados encontrados para cada arquivo de teste relativo aos animais doentes estão mostrados na Tabela 5.1, onde os "Di" e os provenientes de animais sadios são da forma "Si", onde, em ambos, i é o número do arquivo.

Arquivo	Resultado	Arquivo	Resultado
S1	2,08687	D1	0,610853
S2	2,86066	D2	1,31344
S3	1,23819	D3	0,562464
S4	2,00855	D4	0,588281
S5	1,57017	D5	1,2893
S6	1,7743	D6	0,877612
S7	1,69052	D7	1,4017
S8	1,65086	D8	0,0306016
S9	1,84504	D9	1,99726
S10	1,25523	D10	1,39201
S11	1,6944	D11	1,06653
S12	2,05436	D12	0,695742
S13	1,44056	D13	0,902058
S14	0,962914	D14	1,19993
S15	1,49017	D15	0,859135
S16	0,96057	D16	0,737209
S17	0,568798	D17	0,623897
S18	0,314624	D18	1,15054
S19	0,780693		
S20	2,20273		
S21	1,17311		
Esperado:	2	Esperado:	1

Tabela 5.1 - Escala Bark

Arredondando os valores obtidos para se obter um número natural, encontram-se, para os animais doentes, 16 valores como o esperado e, para os animais sadios, 10. Como existem 18 sons de animais doentes e 21 de animais sadios, a porcentagem de acerto foi de 88,9% para os doentes e 47,6% para os sadios.

Destaca-se a quantidade de sons de animais sadios que foram classificados como doentes. Esse fato pode ser explicado, entre outras coisas, pela quantidade de exemplos que foram utilizados para treinar a RNA e pelo fato dos sons serem gravados através do contato de humanos com os animais, o que pode fazer com que animais sadios emitam gritos que se assemelham aos gritos de animais doentes, fazendo com que a RNA se confunda no reconhecimento dos padrões. Para emitirem sons, os animais eram segurados por pessoas, para então ser realizada a gravação. Porém o fato do animal ser segurado por uma pessoa pode induzir a gritos de

insatisfação do animal, independente de estar doente ou não, fazendo com que a rede neural perca muito de sua eficiência em separar os padrões.

Já para os testes com a entalpia e o ITGU, foram utilizados quatro arquivos sonoros, um de cada característica estudada. Desses sons, cada vocalização do animal foi gravada em um arquivo separado, os quais foram processados e geraram os arquivos contendo os vetores de características correspondentes. Este processo foi executado analogamente ao experimento anteriormente descrito. Sendo assim, para a menor entalpia, foram gerados 50 arquivos e para a maior, 133. Já para o menor ITGU, gerou-se 49 arquivos e para o maior, 137. Com isto, foi escolhido separar 45 arquivos de cada característica para treinar a rede neural e utilizar os demais para testá-la.

Iniciando pela entalpia, foram separados 45 sons de menor entalpia. Porém, como existiam mais arquivos de maior entalpia, estes foram separados em três grupos, havendo a intersecção de sons entre eles para que cada grupo possuisse exatos 45 arquivos. Com isto, foram gerados três testes de entalpia, sendo que a entrada de treinamento da rede neural referente à menor entalpia foi sempre os 45 arquivos separados anteriormente e a entrada referente à maior entalpia foi cada um dos grupos supracitados. Para cada grupo de maior entalpia usado no treinamento da rede neural foram utilizados 15 arquivos para testar a rede, sendo os cinco previamente separados de menor entalpia e dez de maior entalpia retirados aleatoriamente dos outros dois grupos que não foram utilizados para treinar a rede. Os resultados estão mostrados nas Figuras 5.1 a 5.4.

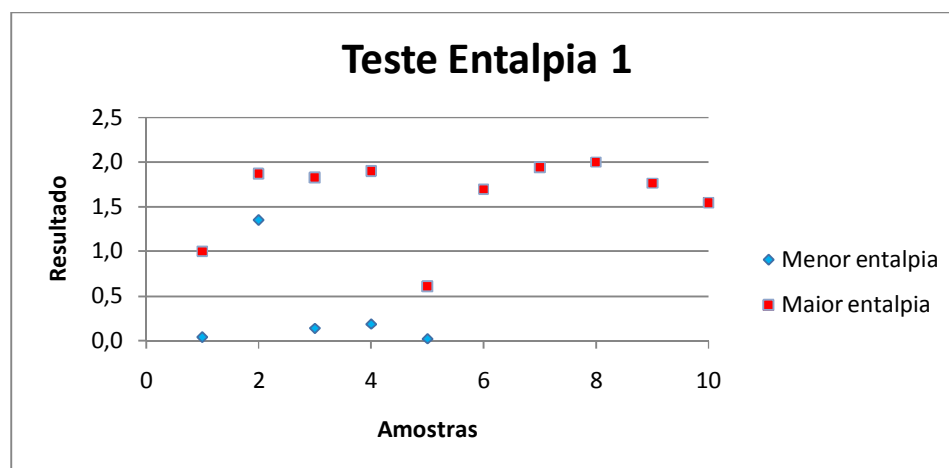


Figura 5.1 - Resultado utilizando o primeiro grupo de maior entalpia

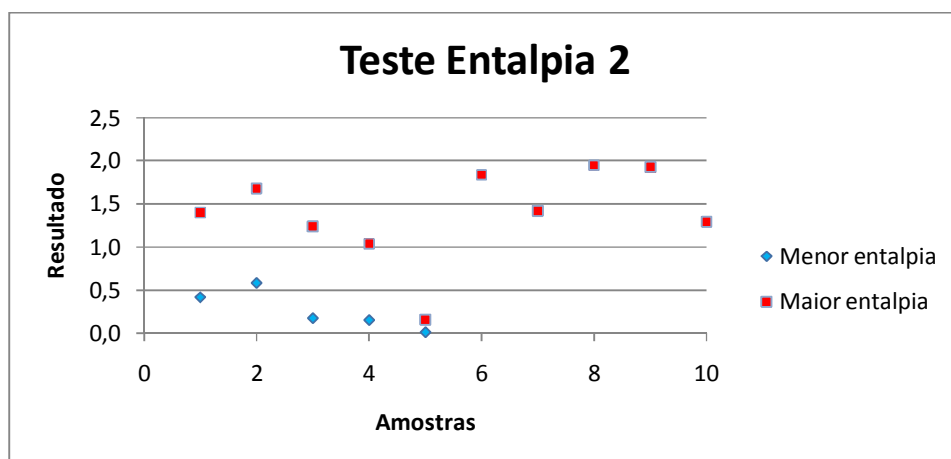


Figura 5.2 - Resultado utilizando o segundo grupo de maior entalpia

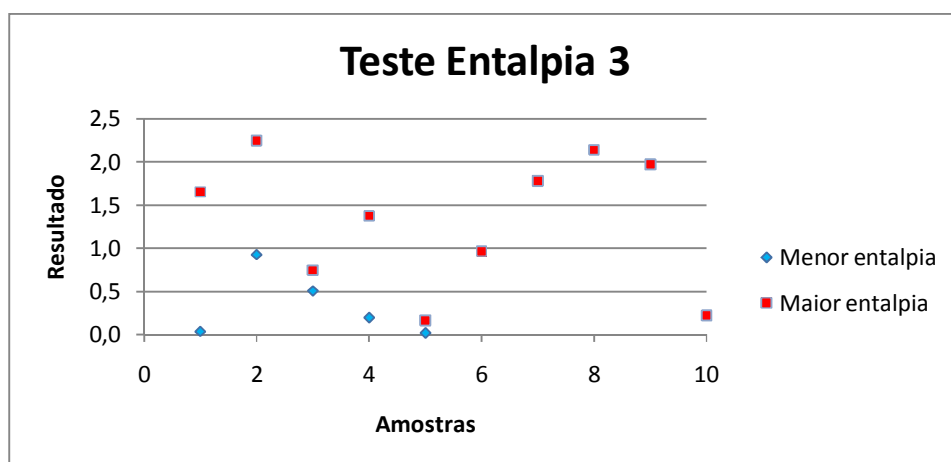


Figura 5.3 - Resultado utilizando o terceiro grupo de maior entalpia

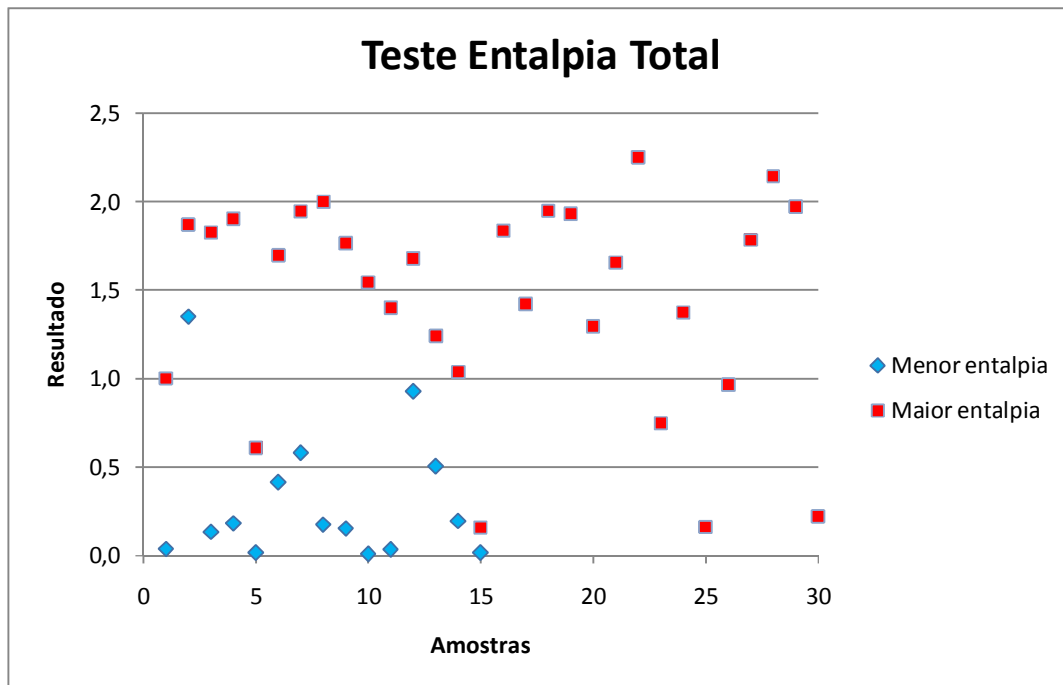


Figura 5.4 - União dos resultados dos teste de entalpia

Analogamente ao procedido com a entalpia, os arquivos sonoros de maior ITGU foram separados em três grupos de 45 arquivos cada um. Nessa separação sobraram dois arquivos, que foram então utilizados para o teste da rede neural. Como dito anteriormente, existiam 49 arquivos sonoros de menor ITGU e, como 45 foram utilizados para treinar a rede, sobraram quatro que foram utilizados para testá-la. Com isso, existiam 16 arquivos de teste para cada grupo utilizado para treinar a rede, sendo 4 provenientes dos sons de menor entalpia, 2 excedentes da separação dos grupos de maior entalpia e 10 dos outros dois grupos que não foram utilizados no teste. Os resultados de cada teste podem ser vistos nas Figuras 5.5 a 5.8.

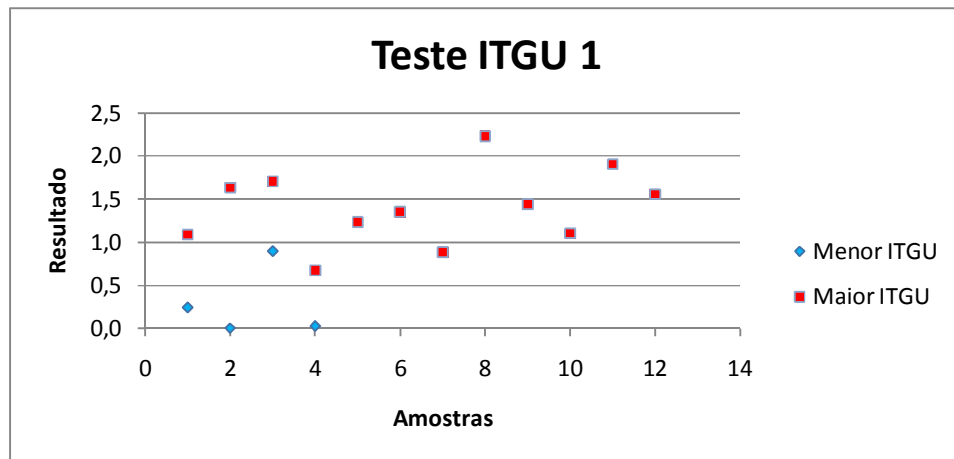


Figura 5.5 - Resultado utilizando o primeiro grupo de maior ITGU

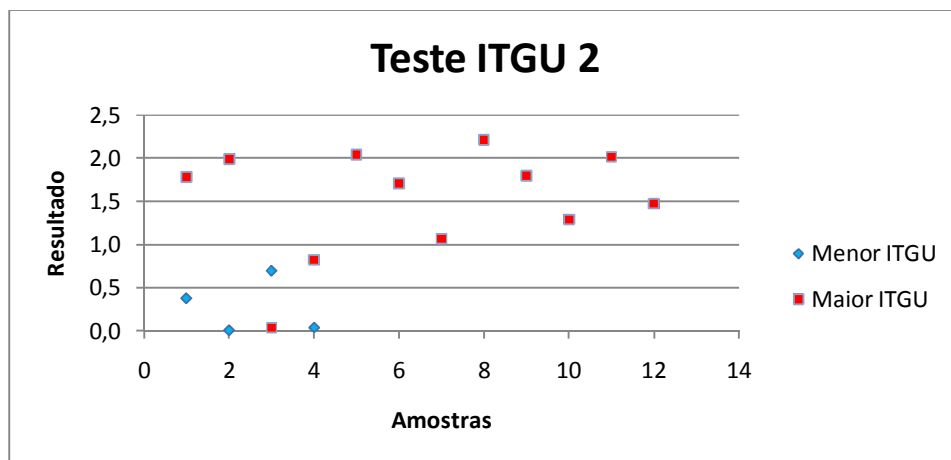


Figura 5.6 - Resultado utilizando o segundo grupo de maior ITGU

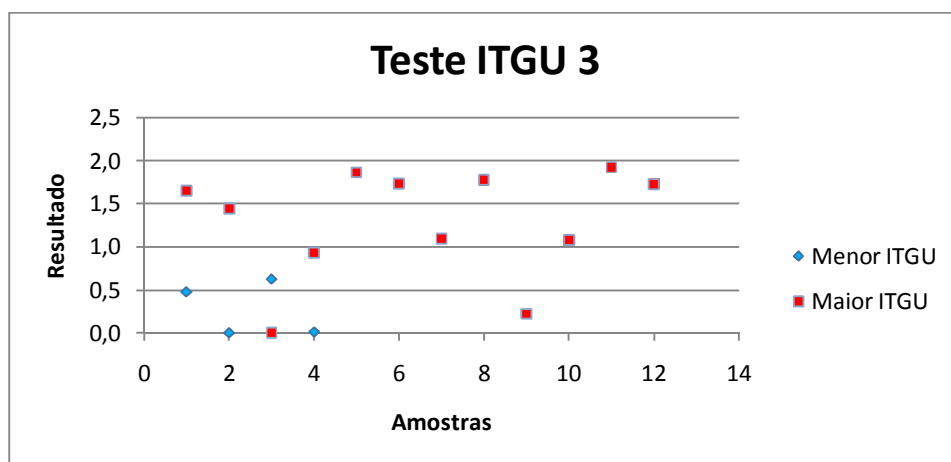


Figura 5.7 - Resultado utilizando o terceiro grupo de maior ITGU

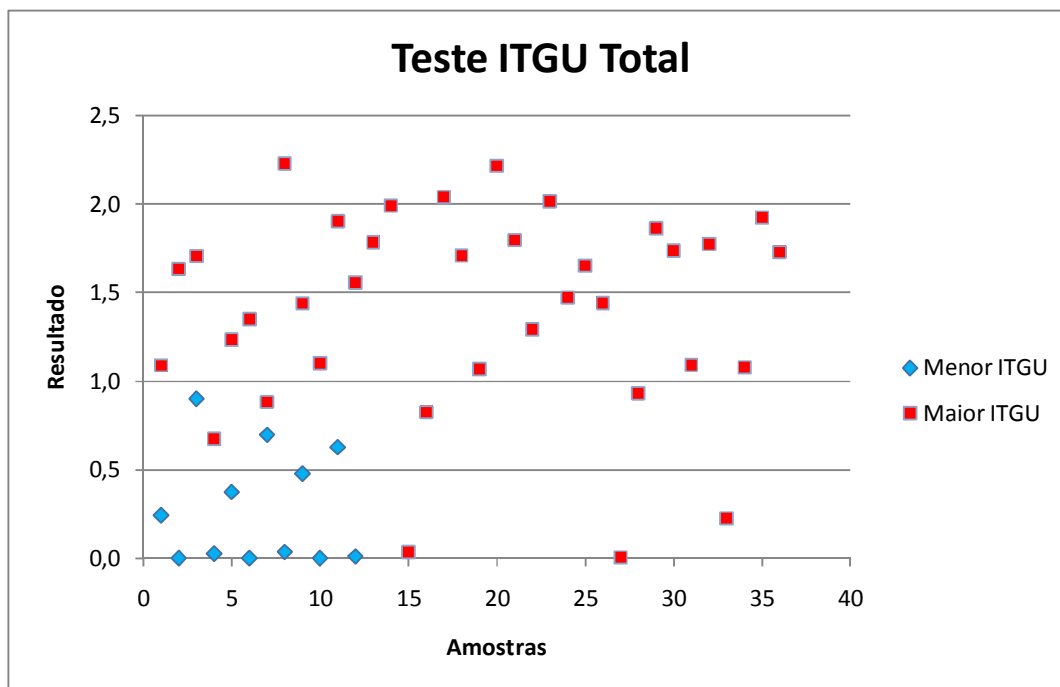


Figura 5.8 - União dos resultados dos testes de ITGU

## 6. Conclusão

Diante das ferramentas utilizadas e citadas, pode-se concluir que a porcentagem de reconhecimento obtida entre os animais doentes foi satisfatória, garantindo a existência de uma semelhança entre os padrões sonoros destes animais. No entanto, para os animais sadios, houve uma baixa porcentagem de reconhecimento, necessitando assim maior aprofundamento nas pesquisas.

Porém, os métodos utilizados mostraram-se eficazes na separação de entalpia e ITGU, uma vez que pode ser vista a existência de um distanciamento entre os resultados obtidos. Contudo, existem alguns pontos que estão fora do resultado geral, gerando alguns falsos positivos. Isto pode ser justificado, mais uma vez, pela precisão da rede neural, que é tão maior quanto mais exemplos forem utilizados em seu treinamento. Em um primeiro momento, poucos exemplos foram utilizados devido à dificuldade existente na gravação dos arquivos sonoros.

Um ponto importante a ser resolvido consiste no treinamento na Rede Neural Artificial, uma vez que este treinamento é muito delicado e passível de erros de interpretação dos padrões. Por isso será necessária a separação de novos arquivos

sonoros contendo os gritos de outros suínos, a fim de se obter uma gama maior de exemplos para a Rede poder reconhecer entre esses padrões de uma forma mais elaborada e precisa.

Outro ponto é a forma de gravação dos sons, que exige contato físico do animal com humanos, que, como já discutido, pode influenciar nas características do grito, confundindo assim a RNA. Logo, seria interessante a obtenção dos gritos dos animais sem interferências humanas, para assim haver uma melhor separação efetiva entre os sons de animais doentes e sadios.

## **7. Referências Bibliográficas**

ALGERS, B. O.; JENSEN P., Communication During Suckling in the Domestic Pig. Effects of Continuous Noise, Applied Animal Behaviour Science, 1985, p. 49–61.

ANIL, K.J.; JIANCHANG, M.; MOHIUDDIN, K.M., Artificial Neural Networks: A Tutorial. in IEEE Computer Society: March 1996; Vol. 29, No. 3 pp. 31-44.

APPLEBY, M.C.; WEARY, D.M.; TAYLOR A. A.; ILLMANN, G., Vocal Communication in Pigs: Who are Nursing Piglets Screaming at?, Ethology 105, 881-892

AUDACITY, HOME PAGE. <<http://audacity.sourceforge.net/?lang=pt>>. Acesso em 01 de novembro de 2010.

BARBEDO, J. G. A., Princípios Fundamentais da Audição Humana, Tese de Doutorado, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, 2004.

BOISSY, A.; MANTEUFFEL, G.; JENSEN, M. B.; OPPERMANN, R.; SPRUIJT, B.; KEELING, L. J. et al, Assessment of positive emotions in animals to improve their welfare, Physiology & Behavior 92, 375-297, 2007.

BOSI, M.; GOLDBERG, R. E., Introduction to Digital Audio Coding and Standards (The Springer International Series in Engineering and Computer Science), Springer, 2002.

BRAGA, A. P., Redes Neurais Artificiais: teoria e aplicações. 1. ed. Rio de Janeiro: Livros Técnicos e Científicos - LTC, 2000. v.1. 250 p.



BRANCO, A.; FEKETE, S. M. W.; RUGOLO, L. M. S. S., REHDER, M. I., Valor e Variações da Frequência Fundamental no Choro de Dor de Recém-Nascidos, Revista CEFAC, São Paulo, v.8, n.4, 529-35, outubro-dezembro, 2006.

FASER, D., The vocalizations and other behaviour of growing pigs in an "open field" test, Applied Animal Ethology, 1, 1974, p. 3-16.

FOURNIER, A., Wavelets and their applications in computer graphics, Siggraph'94, Course Notes, 1995.

GABOR, D., Communication Theory, Past, Present and Prospective, IEEE Trans. Info. Theory, 1953, vol. IT-1, pp. 2-4, February.

GABOR, D., Communication Theory and Physics, IEEE Trans. Info. Theory, 1953, vol. IT-1, pp. 48-59, February.

GASQUET, G.; WITOMSKI, P., Analyse de Fourier et Applications, Paris: Dunod, 2000.

GOMES, J.; VELHO, L.; GOLDENSTEIN, S., Wavelets: Teoria, Software e Aplicações, Rio de Janeiro, IMPA, 216p., 1987.

HILLMANN, E.; MAYER C.; SCHÖN P.C.; PUPPE B.; SCHRADE. L., Vocalisation of domestic pigs (*Sus scrofa domestica*) as an indicator for their adaptation towards ambient temperatures, Applied Animal Behaviour Science 89, 2004, p. 195–206.

JAHNS, G., Call recognition to identify cow conditions - A call-recogniser translating calls to text, Comput. Electron. Agric. (2007), doi:10.1016/j.compag.2007.09.005.

JAIN A. K.; MAO J.; MOHIUDDIN K.M., "Artificial Neural Networks: A Tutorial,"Computer, vol. 29, no. 3, pp. 31-44, Mar. 1996, doi:10.1109/2.485891.

JÚNIOR, S. B., Dynamic Time Warping baseado na Transformada *Wavelet*, Tese de Mestrado, Instituto de Física de São Carlos, Universidade de São Paulo, 2007.

KOVÁCS, Z. L., Redes Neurais Artificiais: Fundamentos e Aplicações, 1996.

LEE, C. H.; CHOU, C. H.; HAN, C. C.; HUANG, R. Z., Automatic recognition of animal vocalizations using averaged MFCC and linear discriminant analysis, Elsevier Science Inc, Nova York, EUA, 2006.

MALLAT, S., Une exploration des signaux en ondelettes, Les Editions de l'École Polytechnique, Palaiseau, 2000.

MANTEUFFEL, G.; PUPPE B.; SCHÖN P. C., Vocalization of farm animals as a measure of welfare, Applied Animal Behaviour Science 88, 2004, p. 163–182.

McCulloch, W. S.; PITTS, W., A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, vol 5, pp 115-133.

MEYER Y.; JAFAR, S.; RIOUL, O., L'Analyse par Ondelettes, Pour La Science, 1987, n. 119, Sept., pp. 28-37.

MORETTIN, P. A., Ondaletas e seus usos em estatística, Mini-curso apresentado na 7ª Escola de Econometria e Séries Temporais, Canela, RS, Brasil, 6-8 Agosto, 1997.

OLIVEIRA, H.M., Análise de Sinais para Engenheiros: uma abordagem via *Wavelets*, 244 pg., Brasport, Rio de Janeiro, 2007

SHEPHERD, G. M.; The synaptic organization of the brain, Oxford University Press, 3<sup>rd</sup> edition, New York, USA, 1990.

SILVA, I. N.; SPATTI, D.H.; FLAUZINO, R.A., Redes Neurais Artificiais para Engenharia e Ciências Aplicadas, 399 pg., Artliber, São Paulo, 2010.

SOTERRONI, A. C.; DOMINGUES, M.O.; RAMOS, F. M., Estimativa do expoente de hurst de séries temporais caóticas por meio da transformada *wavelet* discreta. In: Brazilian Conference on Dynamics, Control and Applications, 7. (DINCON), Presidente Prudente, São Paulo, 2008. Disponível em: <<http://urlib.net/sid.inpe.br/mtc-m18@80/2008/12.18.18.36>>. Acesso em: 01 nov. 2010.

SMITH, S. W., The Scientist and Engineer's Guide to Digital Signal Processing, 1998. Disponível em <<http://www.dspguide.com/pdfbook.htm>>. Acesso em 01 de novembro de 2010.

TAYLOR, A.; WEARY, D.M., Vocal responses of piglets to castration: identifying procedural sources of pain, Applied Animal Behaviour Science 70, 2000, p. 17-26.

WEARY, D.M.; APPLEBY M. C.; FRASER, D., Responses of piglets to early separation from the sow, Applied Animal Behaviour Science 63, 1999, p. 289–300.

ZWICKER, E., Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen), Journal of the Acoustical Society of America 33, pg. 248, 1961.