

UNIVERSIDADE FEDERAL DE GOIÁS
ESCOLA DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA E DE
COMPUTAÇÃO

**Reconhecimento Automático de Aves da Família Tinamidae
Através da Vocalização**

Paulo Francisco da Conceição

Orientador: Prof. Dr. Paulo César Miranda Machado

Goiânia
2012

PAULO FRANCISCO DA CONCEIÇÃO

**Reconhecimento Automático de Aves da Família Tinamidae
Através da Vocalização**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica e de Computação da Escola de Engenharia Elétrica e de Computação da Universidade Federal de Goiás para obtenção do grau de Mestre em Engenharia Elétrica.

Área de Concentração: Engenharia Elétrica

Orientador: Prof. Dr. Paulo César Miranda Machado

Goiânia
2012

PAULO FRANCISCO DA CONCEIÇÃO

Reconhecimento Automático de Aves da Família Tinamidae Através da Vocalização

Dissertação defendida no Programa de Pós-Graduação da Escola de Engenharia Elétrica e de Computação da Universidade Federal de Goiás como requisito parcial para obtenção do título de Mestre no Programa de Mestrado em Engenharia Elétrica e de Computação, aprovada em 28 de Março de 2012, pela Banca Examinadora constituída pelos professores:

Paulo César Miranda Machado, Ph.D

Escola de Engenharia Elétrica e de Computação – UFG
Presidente da Banca

Anamaria Achtschin Ferreira, Dr^a

Ciências Biológicas – UEG

Rodrigo Pinto Lemos, Dr.

Escola de Engenharia Elétrica e de Computação – UFG

Cláudio Afonso Fleury, Dr.

Engenharia Elétrica – UCG

Flávio Henrique Teles Vieira, Dr.

Escola de Engenharia Elétrica e de Computação – UFG

A quem dedicou tempo, recursos e toda a vida por mim, privados de confortos para eu me tornar o que sou hoje — **Meu querido pai e minha querida mãe.**

Agradecimentos

Primeiramente a Deus, Jeová, pela dádiva da vida, força e equilíbrio mental ao lidar com as diversas situações da vida.

“*Se tenho as mãos macias, eu devo tudo a meu pai, que teve as mãos calejadas no tempo que longe vai*” (José Fortuna). Sim, a meu querido pai Lourival Francisco (*in memoriam*) que labutou no limite de sua capacidade, até mesmo física, para me dar sustento e estudo. Este meu querido pai, que me deixou como maior herança, um caráter e valores inigualáveis. Também a minha querida mãe Maria de Fátima, pelo amor, carinho e toda dedicação dada a mim.

Em especial a minha amada esposa Paula e minhas queridas filhas Monique e Lavínia, pela compreensão, pois muito foi o tempo que dediquei a esta pesquisa. Minha esposa querida, tantas vezes tive de deixá-la sozinha... Obrigado mesmo pela sua compreensão e apoio!

Ao meu professor orientador Paulo César, pela credibilidade depositada em mim na realização desta pesquisa e também por todo o tempo dedicado e conhecimento repassado.

Ao professor Rodrigo P. Lemos por todo o tempo dedicado em nortear grande parte da pesquisa, ajustando, por várias vezes, minha linha de pensamento.

Enfim, a todos que de alguma forma contribuíram para que eu pudesse concluir este mestrado, meu muito obrigado.

“A conclusão do assunto, tudo tendo sido ouvido, é: Teme o [verdadeiro] Deus e guarda os seus mandamentos. Pois esta é toda [a obrigação] do homem.”
— O SÁBIO REI SALOMÃO (ECLESIASTES 12:13)

Resumo

O presente trabalho apresenta uma abordagem completa para o desenvolvimento de um sistema de reconhecimento de aves através da vocalização. A abordagem trata especificamente do reconhecimento de aves da família Tinamidae, propondo analisar dados relacionados à frequência e canto da ave, classificando e determinando a espécie da ave. O estudo feito diferencia-se das pesquisas correlatas principalmente por realizar a etapa de pré-processamento de maneira automática, determinando nesta fase as características de frequência mínima, máxima e mais intensa. Ainda é feita uma segmentação do canto da ave, dividindo o mesmo em períodos de som e silêncio, sendo o tempo de canto também usado como característica peculiar de cada espécie analisada. Para a determinação das características de frequência e canto da ave de maneira automática foi feita uma análise da densidade espectral de potência para cada período de tempo especificado em relação à frequência, utilizando para isso o espectrograma do canto. A técnica de classificação e reconhecimento adotada foi a do vizinho mais próximo, aplicando a distância Euclidiana normalizada pelo desvio padrão. A acurácia da técnica utilizada foi de 94,12%.

Palavras-chave: Reconhecimento Automático, Tinamidae, Processamento de Sinais, Tempo-Frequência.

Abstract

This work presents a comprehensive approach to develop a system for recognizing birds by vocalization. The approach specifically addresses the recognition of birds of the Tinamidae family proposing the analysis of data related to the frequency and song of the bird and also classifying and determining the species of bird. The study differs from related research primarily for performing the pre-processing stage automatically. This stage determines the following characteristics: the minimum, the maximum and the stronger frequencies. It's still made a segmentation of the bird singing in periods of sound and silence. The time of singing is also used as a characteristic peculiar to each species analyzed. For the automatic determination of the characteristics of the frequency and song of the bird, an analysis of the power spectral density was made for each time period specified in the frequency using the spectrogram of the song. The recognition and classification technique adopted was the nearest neighbor, using Euclidean distance normalized by the standard deviation. The accuracy of the technique used was 94.12%.

Keywords: Automatic Recognition, Tinamidae, Signal Processing, Time-Frequency.

Sumário

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas

1	Introdução	14
1.1	Breve Histórico Sobre as Aves	14
1.2	Classificação Taxonômica das Aves	16
1.3	Vocalização das Aves	18
1.4	O Cerrado e a Família Tinamidae	20
1.5	Pesquisa Realizada	21
2	Metodologia	23
2.1	Introdução	23
2.2	Sinal Discreto	23
2.2.1	Domínio do Tempo	24
2.2.2	Domínio da Frequência	25
2.3	Padronização dos Sinais	26
2.3.1	Formato do Arquivo	27
2.3.2	Normalização e Taxa de Amostragem	27

2.4	Filtros Digitais	28
2.4.1	Filtro <i>Butterworth</i>	30
2.4.2	Filtro <i>Chebyshev</i>	31
2.4.3	Filtro <i>Chebyshev</i> Invertido	31
2.4.4	Filtro Elíptico	32
2.4.5	Escolha do Filtro	33
2.5	Interferência de 60 Hz - Filtro Pente	34
2.6	Detecção Automática de Frequências	36
2.7	Estimativa do Tempo de Canto	37
2.8	Método de Reconhecimento – Vizinho Mais Próximo	38
2.9	Conclusão	40
3	Experimentos e Resultados	42
3.1	Introdução	42
3.2	Pré-processamento	45
3.2.1	Detectando e Eliminando Interferência de 60 Hz	45
3.2.2	Determinação de Frequências	46
3.2.3	Aplicação do Filtro Passa-banda	47
3.2.4	Determinação do Tempo de Canto	49
3.3	Treinamento	50
3.4	Reconhecimento	53
3.5	Conclusão	57
4	Conclusão e Trabalhos Futuros	58
	Referências	60
	Anexo A - Lista de Artigos Publicados	64

Lista de Figuras

Figura 1 - Distribuição da quantidade de espécies em alguns países da América. Fonte: Avibase (2011).	15
Figura 2 - Síntese da taxonomia das aves.	16
Figura 3 - Espectrograma do canto de uma Perdiz – <i>Rhynchotus rufescens</i>	19
Figura 4 - Perdiz - <i>Rhynchotus rufescens</i>	20
Figura 5 - Diagrama de blocos mostrando a sequência do pré-processamento proposto neste trabalho.	24
Figura 6 - Exemplo de sinal no domínio do tempo.	25
Figura 7 - Exemplo de sinal no domínio da frequência.	26
Figura 8 - Resposta da aplicação dos filtros (a) Passa-baixas, (b) Passa-altas, (c) Passa-banda e (d) Rejeita-banda. (Adaptada de Smith (1999, p. 268)).	29
Figura 9 - Parâmetros que avaliam o desempenho de um filtro passa-baixas.	30
Figura 10 - Resposta do filtro <i>Butterworth</i>	31
Figura 11 - Resposta do filtro <i>Chebyshev</i>	32
Figura 12 - Resposta do filtro <i>Chebyshev</i> Invertido.	32
Figura 13 - Resposta do filtro Elíptico.	33
Figura 14 - Resultado do processo de filtragem para a vocalização de um Azulona.	35
Figura 15 - Resposta do filtro Pente.	36
Figura 16 - Exemplo da detecção de regiões de som para o canto de uma Perdiz, sendo detectadas duas regiões de som de igual tamanho.	38

Figura 17 - Sinal e espectro de uma Perdiz, evidenciando a interferência de 60 Hz e seus harmônicos.	45
Figura 18 - Sinal e espectro de uma Perdiz, evidenciando a clareza do sinal após a aplicação do filtro pente, que eliminou a interferência de 60 Hz e seus harmônicos.	46
Figura 19 - Sinal e espectro de uma Perdiz, evidenciando a inspeção visual do espectrograma e constatação da faixa de frequência.	47
Figura 20 - Sinal e espectro de uma Perdiz, filtrado com parâmetros determinados automaticamente.	49
Figura 21 - Inspeção visual para detecção da região de som do sinal de um Jaó. . . .	50
Figura 22 - Detecção automática para encontrar a região de som do sinal de um Jaó. . . .	50
Figura 23 - Disposição gráfica dos parâmetros coletados por espécie - Tempo x Frequência.	54
Figura 24 - Sequência do processo de reconhecimento de um sinal de entrada. . . .	54
Figura 25 - Determinação das características de um sinal de entrada.	55
Figura 26 - Disposição gráfica com sinal de entrada.	56

Lista de Tabelas

Tabela 1 -	Listagem de algumas ordens e famílias da avifauna brasileira.	17
Tabela 2 -	Espécies da Família Tinamidae Estudadas	43
Tabela 3 -	Comparação entre inspeção (IVis) visual e detecção automática (DAut) das frequências mínimas (FMín), mais intensas (FInt) e máximas (FMáx).	48
Tabela 4 -	Comparação entre inspeção visual (IVis) e detecção automática (DAut) do tempo de canto (MédC).	51
Tabela 5 -	Dados coletados de <i>Crypturellus undulatus</i> (Jaó).	51
Tabela 6 -	Dados coletados de <i>Crypturellus strigulosus</i> (Inhambu relógio).	52
Tabela 7 -	Dados coletados de <i>Rhynchotus rufescens</i> (Perdiz).	52
Tabela 8 -	Tabela de treinamento para as espécies estudadas.	53
Tabela 9 -	Distâncias calculadas para o exemplo de um <i>Crypturellus strigulosus</i>	55
Tabela 10 -	Resultados obtidos na classificação de algumas espécies.	56
Tabela 11 -	Matriz de confusão dos sinais testados.	57

Lista de Abreviaturas

<i>dB</i>	Decibel
<i>DSP</i>	Digital Signal Processing (Processamento Digital de Sinais)
<i>DTFT</i>	Discret-Time Fourier Transform (Transformada de Fourier para Sistemas Discretos)
<i>FFT</i>	Fast Fourier Transform (Transformada Rápida de Fourier)
<i>GMM</i>	Gaussian Mixture Model (Modelo de Misturas Gaussianas)
<i>HMM</i>	Hidden Markov Model (Cadeias Ocultas de Markov)
<i>Hz</i>	Hertz
<i>INPA</i>	Instituto Nacional de Pesquisas na Amazônia
<i>MATLAB</i>	MaTriX Laboratory
<i>mp3</i>	MPEG 1 Layer-3
<i>PSD</i>	Power Spectral Density (Densidade Espectral de Potência)
<i>RNA</i>	Redes Neurais Artificiais
<i>seg</i>	Segundos
<i>wav</i>	WAVE form audio format

Introdução

As aves há muito tempo são objetos de desejo, apreciação e estudo. No decorrer da história, a humanidade passou a admirar estes animais alados, chegando algumas pessoas a ponto de prender aves em gaiolas, tendo-os sempre por perto para admirar sua beleza e seu canto (MARQUES, 2009 apud SICK, 1997, p. 58). Há ainda outros, como Save (2010), que afirmam categoricamente que o lugar de toda espécie de aves é na natureza, sendo deste modo a melhor forma de respeitá-las e admirá-las. Save (2010), ainda ressalta que as aves podem ser consideradas indicadores do bem estar do homem e do meio ambiente, de tal forma que o que acontece com ecossistema reflete nas as aves.

A vocalização das aves é um dos objetos de estudo desta pesquisa, que tem como objetivo final, utilizando técnicas de Processamento Digital de Sinais, analisar o sinal sonoro das aves para realizar a classificação/reconhecimento automático da espécie à qual a ave pertence. Embora as técnicas estudadas, de maneira geral, possam ser usadas para classificar aves de diversas famílias, concentrou-se as pesquisas em procurar reconhecer aves da família dos tinamídeos.

Para melhor entender os motivos que levaram a esta escolha, bem como as técnicas que foram adotadas para cumprir o objetivo proposto, primeiramente serão descritas algumas características relevantes das aves de modo geral e da família Tinamidae.

1.1 Breve Histórico Sobre as Aves

Morfológicamente as aves se caracterizam por serem animais vertebrados com membros anteriores modificados para voo, corpo revestido de penas e um bico córneo, desprovido de dentes (AULETE, 2011). Estes vertebrados habitam todos os ecossistemas da Terra, podendo ser encontrados desde o clima mais quente até o clima mais frio. Estima-se que em todo o mundo exista cerca de 10.000 espécies de aves (BIRDLIFE, 2011).

Portanto, dentre todos animais da Terra, nenhum soube compartilhar e usufruir tão bem dos recursos da natureza quanto as aves. Descourtilz (1983), até mesmo diz, poeticamente, que elas foram “distribuídas por tribos diferentes ao redor do planeta, não existindo nenhuma zona do mundo em que o homem não as encontre”.

É fato que para a maioria das aves sobreviverem é necessário um ambiente favorável. Sendo assim, por causa da devastação do planeta e também pela caça descontrolada, muitas espécies de aves foram extintas e ainda outras estão seriamente ameaçadas de extinção (SOUZA, 1987).

Mesmo com este fato alarmante muitos não se importam, pois segundo Marques (2009), há ainda, entre muitas pessoas, uma cultura arraigada de manter aves em cativeiro, principalmente os pássaros canoros. Esta prática contribui bastante para a diminuição de várias espécies de aves, podendo chegar ao ponto de serem extintas.

Entretanto, nota-se que ao passar de cada ano o número de espécies tende a aumentar, isto porque ainda não haviam sido identificadas ou classificadas. No que se refere ao Brasil, de acordo com CBRO (2010), existem atualmente cerca de 1832 espécies de aves registradas contra 1590 no ano de 1981, o que sempre deixou o Brasil entre os países com o maior número de espécies do mundo.

A figura 1 mostra a distribuição das espécies de aves em alguns dos principais países da América. A Colômbia possui a maior avifauna do mundo e é possível perceber que o Brasil está em segundo lugar e ainda possui um grande número de espécies endêmicas¹, cerca de 216. Entretanto o Brasil detém o maior número de espécies em extinção, totalizando 120 espécies (SAVE, 2010).

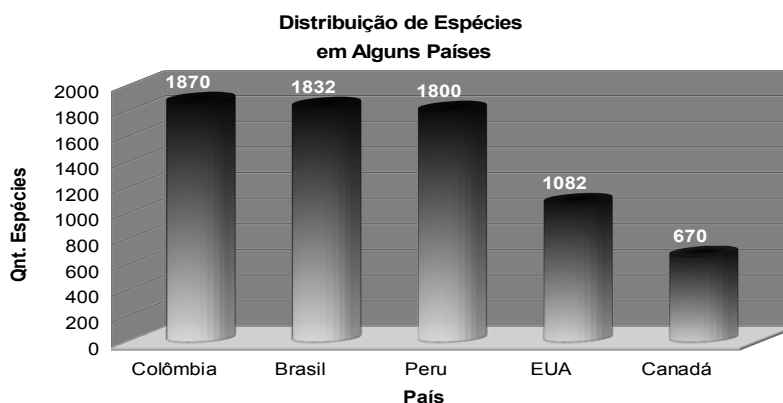


Figura 1 – Distribuição da quantidade de espécies em alguns países da América.

Fonte: Avibase (2011).

¹ Espécie que é restrita a determinada região ou área geográfica específica (BROWN; LOMOLINO, 2006).

1.2 Classificação Taxonômica das Aves

De acordo com com Antas e Cavalcanti (1988), as espécies de aves têm parentesco entre si, com características comuns dentro de cada grupo. Assim os ornitólogos classificam as aves baseados nessas características comuns. Sick (1984), ressalta que para identificar e classificar cientificamente uma ave, são necessárias várias análises e medições exatas da coloração, medida das asas, plumagem e outras características.

Sick (1984), ainda diz que em muitos casos as espécies são morfologicamente tão semelhantes que apenas o canto é diferente, tomando então este como base para realizar a classificação da espécie.

O estudo da classificação científica dos animais é campo da taxonomia ou zoologia sistemática. Atualmente uma das taxonomias mais utilizadas é a proposta em 1756 pelo cientista sueco Carl Von Lineu. A nomenclatura científica é toda em termos latinizados e a classificação, como já dito, espelha o grau de parentesco e similaridade, da forma como é analisada pelos cientistas. Assim algumas aves podem estar em famílias diferentes, mas pertencerem a mesma ordem (ANTAS; CAVALCANTI, 1988).

Desta forma, em todo o reino animal (Animalia), a classe dos vertebrados que possuem penas é denominada Aves. Daí então, basicamente, a taxonomia das espécies tem as categorias mostradas na figura 2, podendo existir, entretanto, algumas subclassificações das que são apresentadas (CBRO, 2010).

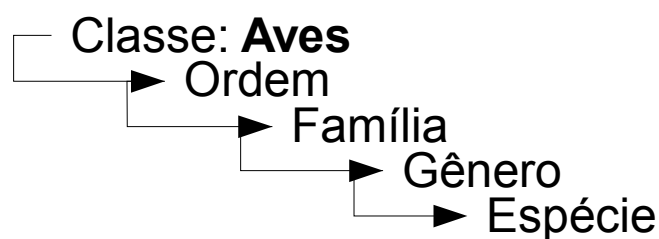


Figura 2 – Síntese da taxonomia das aves.

De acordo com os dados de Avibase (2011) existem, na classe das aves, aproximadamente 59 ordens e subordens, totalizando cerca de 230 famílias, considerando ordens e famílias já extintas. Já de acordo com CBRO (2010), no Brasil existem, no início de 2011, cerca de 1.832 espécies, espalhadas em 56 famílias divididas em 31 ordens. Algumas ordens e famílias da avifauna brasileira são listadas na tabela 1.

É interessante ressaltar que no Brasil, assim como de modo geral, a maioria das aves per-

Tabela 1 – Listagem de algumas ordens e famílias da avifauna brasileira.

Classe	Ordem	Família
Aves	Struthioniformes	Rheidae
	Tinamiformes	Tinamidae
	Anseriformes	Anhimidae
		Anatidae
	Galliformes	Cracidae
		Odontophoridae
	Podicipediformes	Podicipedidae
	Sphenisciformes	Spheniscidae
	Procellariiformes	Diomedidae
		Procellariidae
	Phaethontiformes	Phaethontidae
	Ciconiiformes	Ciconiidae
	Suliformes	Fregatidae
		Sulidae
	Pelecaniformes	Ardeidae
		Threskiornithidae
	Phoenicopteriformes	Phoenicopteridae
	Cathartiformes	Cathartidae
	Accipitriformes	Pandionidae
		Accipitridae
	Falconiformes	Falconidae
	Eurypygiiformes	Eurypygidae
	Gruiformes	Psophiidae
		Rallidae
	Cariamiformes	Cariamidae
	Charadriiformes	Scolopacidae
		Sternidae
	Columbiformes	Columbidae
	Psittaciformes	Psittacidae
	Opisthocomiformes	Opisthocomidae
	Cuculiformes	Cuculidae
	Strigiformes	Tytonidae
		Strigidae
Caprimulgiformes	Steatornithidae	
	Nyctibiidae	
Apodiformes	Apodidae	
	Trochilidae	
Trogoniformes	Trogonidae	
Passeriformes	Thamnophilidae	
	Tyrannidae	
	Furnariidae	
	Thraupidae	
	Emberizidae	
	Rhynchocyclidae	
	Dendrocolaptidae	
	Passeridae	
	Icteridae	
Pipridae		

Fonte: Avibase (2011)

tence à ordem dos Passeriformes. As aves desta ordem são popularmente conhecidas por pássaros canoros, pois os sons destas aves são melódiosos. No mundo existem cerca de 5.400 espécies de pássaros, o que é mais da metade de todo o universo das espécies. No Brasil, segundo o último registro em CBRO (2010), existem cerca de 1023 espécies de pássaros, o que também é mais da metade das espécies brasileiras.

1.3 Vocalização das Aves

Uma das formas mais práticas de se identificar uma ave é através do som emitido por esta, tendo em vista que de modo geral, cada espécie de ave produz um som diferente. Segundo o observador Helmut Sick, 90% dos levantamentos de aves durante uma observação acontece pelo canto, pois em uma mata densa muitas vezes não é possível enxergar o pássaro (SICK, 1984).

No intuito de analisar mais profundamente a vocalização dos animais em geral, surgiu uma ciência denominada bioacústica que, no caso das aves, tem revelado dados importantes, como a identificação de espécies e motivação da vocalização das aves (MARQUES, 2008). Embora Sick (1984), afirmasse que a análise da vocalização das aves pudesse apenas auxiliar no reconhecimento de uma espécie, estudos e testes recentes, como os indicados em Correio (2010), levam a crer que quando o som de uma ave é diferente de outra, isto indica que, mesmo sendo aparentemente semelhantes, são morfologicamente diferentes, indicando assim ser de outra espécie.

De acordo com Marques (2008), as aves produzem sons para os mais variados fins como acasalamento, territorialidade e também comunicações de sinais de alerta. Além disso, Marques (2008), acrescenta que a vocalização da ave dá sinais sobre o seu temperamento e habilidades.

Aves produzem sons principalmente através da siringe e têm limites de audição bem parecidos com os do ser humano, entre 20 Hz e 20.000 Hz. De acordo com Sick (1984), as aves podem produzir sons muito graves, na faixa da frequência de 80 Hz e também sons muito agudos, na faixa de 12.000 Hz, chegando seus ultrassons até 18.000 Hz. De acordo com Dias (2009), Krause (1993), a faixa de frequência da vocalização das aves está ligada também ao ambiente em que a ave está. A frequência da voz de muitos passeriformes é de 4.000 Hz, o que corresponde a frequência aproximada da tecla mais alta de um piano (SICK, 1984).

Sick (1984), afirma que "a maioria das aves, se distingue por uma chamada, um grito ou um pio que emite em várias ocasiões". Estes chamados podem ser emitidos, por exemplo, quando a ave vai voar ou simplesmente para identificação individual.

Para Sick (1984), deve ficar claro que chamado é diferente de canto, embora ainda confesse

faltar um limite entre as definições. Para ele, canto pode ser entendido como sendo uma simples repetição de uma mesma nota ou os melódiosos e harmônicos sons emitidos pelos passeriformes.

De acordo com Marques (2009), as aves podem mudar a variação do seu canto de acordo com o ambiente em que vivem. Além do mais a intensidade do canto varia muito dependendo do ambiente em que as aves de modo geral estão: nas cidades onde há muita presença de ruído, os sons tendem a ser mais altos, enquanto nas florestas e matas, os sons embora se propaguem mais, tem menos intensidade (SICK, 1984).

Atualmente os sonogramas, ou espectrogramas, permitem analisar nitidamente a manifestação sonora das aves, tornando possível identificar a fonética de cada canto que, de acordo com Somervuo *et al.* (2006) e Fagerlund (2004) pode, geralmente, ser dividido da seguinte forma hierárquica:

- **Nota** – É o menor elemento de som, sem intervalo, presente no canto;
- **Sílaba** – Conjunto de uma ou mais notas. Há um pequeno intervalo entre uma sílaba e outra e normalmente elas tem o mesmo padrão;
- **Frase** – Conjunto de uma ou mais Sílabas. Quando o padrão de sílabas é alterado ou quando há um grande intervalo entre uma sílaba e outra, diz-se que a frase terminou e iniciou outra.

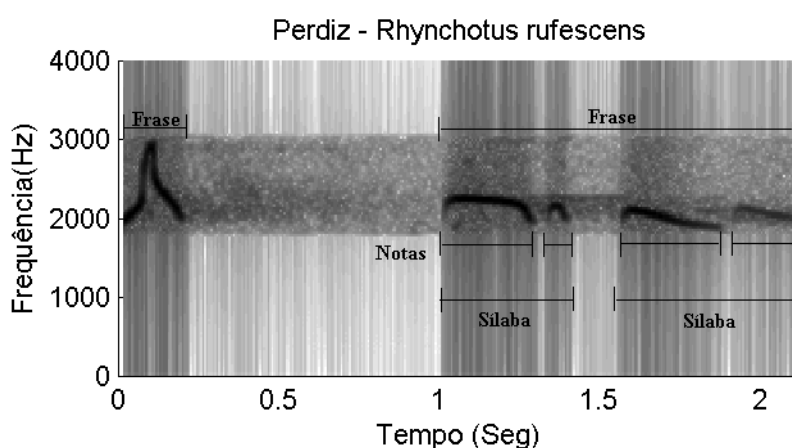


Figura 3 – Espectrograma do canto de uma Perdiz – *Rhynchotus rufescens*.

Na figura 3 é mostrado o espectrograma para um canto completo de uma Perdiz. É possível perceber a variação da intensidade dos componentes de frequência, presentes no canto, ao longo do tempo. As partes mais escuras são do canto da ave. É possível ainda perceber as pausas ao longo do tempo, evidenciando assim as sílabas e frases que constituem o canto completo da ave.

No caso dos tinamídeos, por terem um canto geralmente breve e com poucos intervalos, normalmente sua vocalização é composta por poucas frases, sílabas e notas, diferentemente dos pássaros canoros, em que esta manifestação fica mais evidente.

1.4 O Cerrado e a Família Tinamidae

Segundo Silva e Santos (2005), o Cerrado é um dos biomas mais ricos da avifauna brasileira, com cerca de 856 espécies, das quais 30 são endêmicas e 14 estão ameaçadas. Entretanto, o cerrado tende a diminuir rapidamente, uma vez que seus ambientes naturais estão cada vez mais fragmentados e são destruídos de forma rápida, sendo substituídos pela agricultura e pecuária. Ainda assim, de acordo com Felfili *et al.* (2005), o cerrado ocupa 24% do território brasileiro, entretanto, apenas algumas áreas do cerrado são legalmente protegidas, o que contribui mais ainda para a destruição deste ambiente.

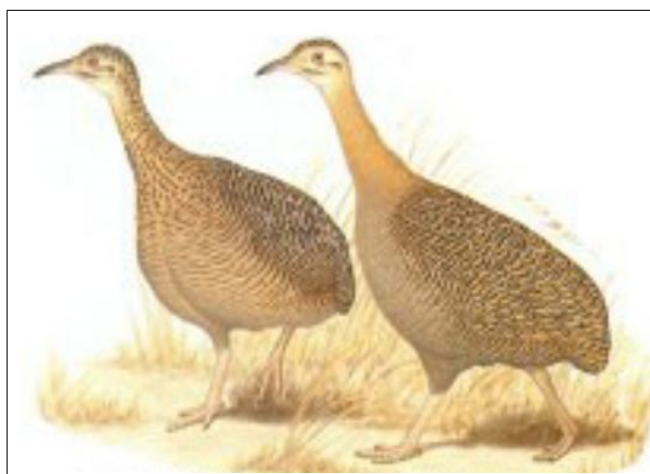


Figura 4 – Perdiz - *Rhynchotus rufescens*.

Este rico e ameaçado ambiente da avifauna é também moradia de algumas espécies da família dos tinamídeos (Tinamidae), que é composta por 23 espécies, das quais mais de 90% vivem no Brasil em ambientes como o cerrado (WIKIAVES, 2011b). São aves tipicamente terrestres e, como mostra a figura 4, são de aspecto galináceo apresentando cores miméticas no colorido da plumagem, em adaptação aos seus hábitos terrestres. Esta família de aves, além de sofrerem com a degradação do ambiente em que vivem, ainda sofrem com caçadores, pois a carne destas aves é considerada saborosa (WIKIAVES, 2011a; SILVEIRA, 2010).

Com relação a vocalização dos tinamídeos, Silveira (2010, p. 68) ressalta que uma das “características mais marcantes dos tinamídeos é a beleza e a qualidade do canto. Sua vocalização varia entre sons aflautados, doces e simples até trinados altos, que podem ser ouvidos a longas

distâncias". No intuito de definir uma faixa de frequência da vocalização dos tinamídeos, foram feitos testes com todas as 23 espécies da família e concluiu-se que podem emitir sons que variam aproximadamente entre 300 Hz e 4.000 Hz.

1.5 Pesquisa Realizada

O monitoramento de espécies de aves é de suma importância, isto porque a avifauna tem sido usada nos estudos da qualidade dos ambientes e as aves ainda são reconhecidas como bioindicadoras dos ecossistemas terrestres (MORRISON, 1986; ALMEIDA; ALMEIDA, 1998; ALMEIDA, 1981; ALEGRINI, 1997). Além disso o monitoramento e identificação de espécies pode contribuir para a conservação das aves, em especial no cerrado que, de acordo com Marini e Garcia (2005), possui o segundo maior número de espécies ameaçadas. Por causa da devastação, Pinheiro (2007), destaca que o Cerrado tem se tornado um ambiente fragmentado, o que para Almeida e Almeida (1998) favorece o empobrecimento das espécies.

Tem havido recentemente grande interesse nas pesquisas sobre avaliação computacional dos sons de aves para diversos fins, como identificação comportamental e detecção de espécies (KWAN *et al.*, 2004; CORREIO, 2010; SELIN *et al.*, 2007; LEE *et al.*, 2008; BRANDES, 2008).

Kwan *et al.* (2004), em sua pesquisa, realizam a identificação e classificação de espécies através do Modelo de Misturas Gaussianas (GMM). Já Selin *et al.* (2007) exploram a utilização de *wavelets*, enquanto Lee *et al.* (2008) utilizam coeficientes cepstrais para o reconhecimento automático de espécies.

O resultado da maioria das pesquisas relacionadas a técnicas de classificação e reconhecimento automático de aves, como indicado em Correio (2010), é a possibilidade de catalogação e até mesmo a descoberta de novas espécies, tendo em vista que com análise da vocalização é possível identificar características únicas para cada espécie. Com a catalogação e monitoração de espécies é possível contribuir com dados estatísticos que podem auxiliar na preservação das mesmas.

Para exemplificar, Correio (2010), menciona que de acordo com o Instituto Nacional de Pesquisas na Amazônia - INPA, acredita-se que o número de espécies de aves da Amazônia seja pelo menos o dobro das cerca de 1.300 espécies anteriormente registradas. A identificação destas novas espécies pode se dar de forma rápida e prática devido a aplicação de técnicas de análise da vocalização das aves.

Assim, acredita-se que técnicas de reconhecimento rápido das aves podem contribuir para

preservação e conservação da avifauna, podendo catalogar as aves que estão migrando para as cidades e até identificar novas espécies (CORREIO, 2010; SELIN *et al.*, 2007).

Conforme já mencionado, o cerrado é um ambiente que sofre degradação bem acentuada e possui um grande número de espécies, incluindo os tinamídeos, assim monitorar as aves desta região de uma forma rápida e eficaz é importante (MARINI; GARCIA, 2005).

Pensando nisso é proposto nesta pesquisa analisar técnicas de Processamento Digital de Sinais – DSP (Digital Signal Processing), com a finalidade de tratar os sinais sonoros da vocalização das aves da família dos tinamídeos, no intuito de extrair características do sinal que possam servir para uma classificação e identificação automática da espécie a qual a ave pertence.

Para aplicar as técnicas estudadas, foram usados sons de aves da família dos tinamídeos, por se tratarem de aves tipicamente brasileiras, sendo que algumas habitam a região do cerrado, e produzirem sons bem característicos (WIKIAVES, 2011b).

No capítulo 2 serão apresentadas as operações aplicadas ao sinal de entrada para realizar o pré-processamento, destacando os filtros aplicados e metodologia para extração de frequências e de segmentação do sinal em partes distintas de som e silêncio. Ainda neste capítulo será apresentado o método estatístico do vizinho mais próximo, como técnica de treinamento e reconhecimento.

No capítulo 3 serão mostrados e discutidos os resultados obtidos para cada etapa do reconhecimento: pré-processamento, treinamento e classificação. Serão mostrados gráficos e tabelas, indicando a acurácia da abordagem sugerida na realização do reconhecimento para 11 espécies de aves da família Tinamidae.

Metodologia

2.1 Introdução

Neste capítulo serão apresentadas as técnicas propostas para preparação do sinal ¹ que antecedem a etapa de classificação e reconhecimento, conforme sequência mostrada na figura 5.

A realização do pré-processamento é necessária, pois na maioria das vezes em que o sinal é coletado, vários fatores fazem com que ele seja corrompido. Dentre os possíveis distúrbios causados ao sinal original, podem ser citados: interferência da rede elétrica, barulhos e sons do ambiente de coleta. Assim, qualquer som que não seja da ave a ser analisada pode ser considerado um ruído, que pode interferir diretamente no acerto da classificação e reconhecimento, sendo necessário, portanto, atenuar este(s) ruído(s).

Também na etapa de pré-processamento serão determinadas características relacionadas às frequências, que servirão como parâmetros do classificador. Estas características tem a ver com a frequência mínima, máxima e mais intensa do canto da ave. Ainda nesta fase o canto da ave é segmentado em períodos de som e silêncio, determinando assim mais parâmetros de classificação.

Antes, porém, será mostrado como um sinal de tempo discreto pode ser representado e compreendido no domínio do tempo e no domínio da frequência.

2.2 Sinal Discreto

Um sinal analógico é assim chamado porque varia sua intensidade continuamente ao longo do tempo em analogia à grandeza física que representa. Entretanto, caso este sinal seja limitado

¹ Desde que não esteja especificado, ao usar a palavra sinal, este texto estará se referindo a sinal de tempo discreto correspondente ao canto de aves.

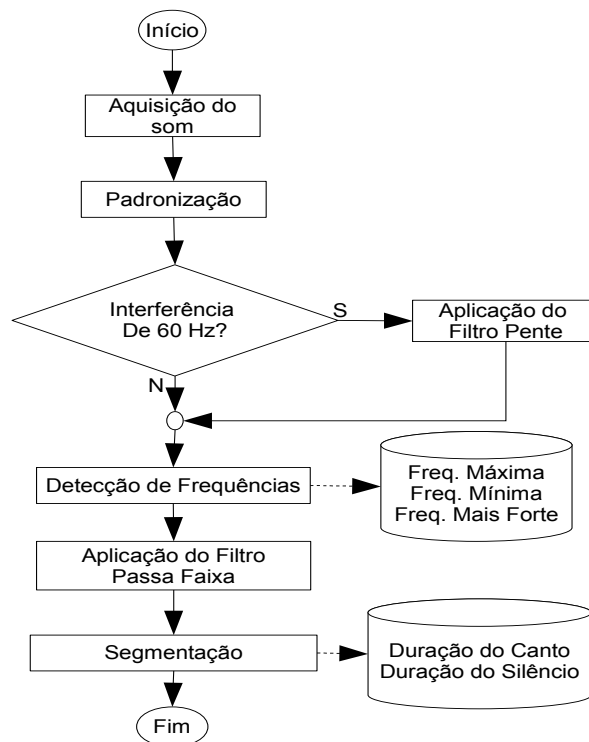


Figura 5 – Diagrama de blocos mostrando a sequência do pré-processamento proposto neste trabalho.

a uma frequência, pode ser representado unicamente a partir de amostras tomadas a intervalos de tempo regulares. Adicionalmente, para possibilitar o processamento digital, é necessário promover também a discretização e codificação binária destas amostras.

Segundo Oppenheim e Schafer (1999), o sinal digital pode ser representado matematicamente por uma lista de valores ordenados. Esta sequência numérica é definida formalmente como

$$x = \{x[n]\}, -\infty < n < \infty, \quad (2.1)$$

em que n é um número inteiro. Assim, x representa todo o sinal amostrado, enquanto cada $x[n]$ representa um valor no instante n de tempo.

Tendo os valores de x , é possível efetuar as técnicas de Processamento Digital de Sinais (DSP). A representação do sinal bem como seu processamento, podem ser feitos tanto no domínio do tempo como no domínio da frequência.

2.2.1 Domínio do Tempo

Segundo Haykin e Veen (2001), no domínio do tempo a amostra é distribuída de maneira igualmente espaçada, podendo ser mostrada a amplitude do sinal distribuída ao longo do tempo.

A figura 6 ilustra o aspecto de um sinal do canto de uma ave visualizado no domínio do tempo², evidenciando no eixo horizontal o tempo em segundos e no eixo vertical a amplitude. Nota-se que, o sinal mostrado na figura 6, tem aproximadamente 1,6 segundos de duração.

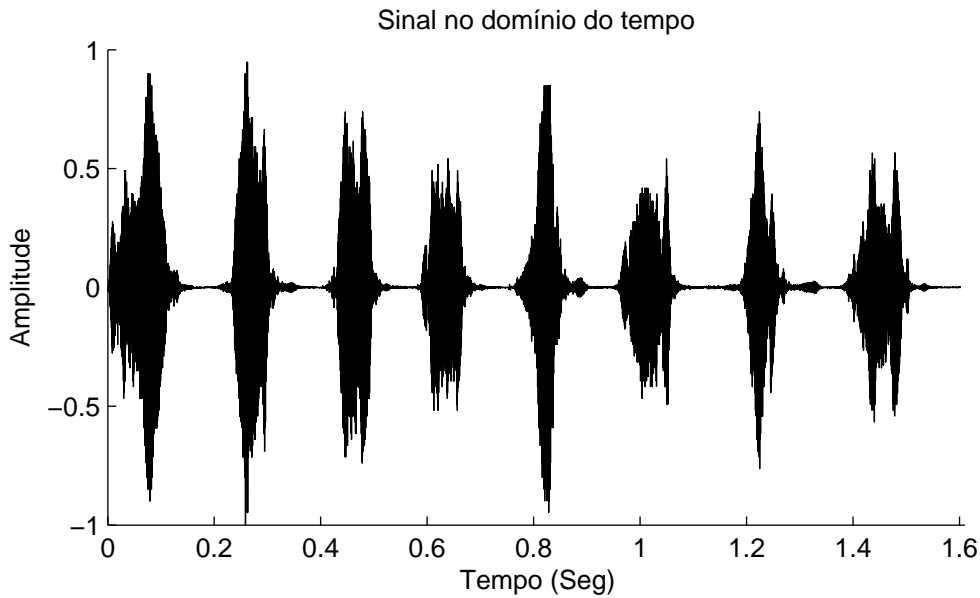


Figura 6 – Exemplo de sinal no domínio do tempo.

2.2.2 Domínio da Frequência

Com a representação do sinal no domínio da frequência, também chamado de espectro do sinal, é possível a realização de diversas operações, como a aplicação de filtros digitais e extração de características relacionadas às frequências (OPPENHEIM; SCHAFER, 1999).

Esta representação do sinal no domínio da frequência pode ser obtida com cálculo da Transformada de Fourier para Sinais Discretos (DTFT), segundo Oppenheim e Schafer (1999) definida como

$$X[k] = \sum_{n=1}^{N-1} x[n]e^{-i\omega n}, \quad (2.2)$$

onde $X[k]$ é o espectro em frequência obtido da decomposição do sinal de entrada x para cada instante de tempo n em uma base de funções exponenciais complexas $e^{-i\omega n}$, sendo N o número total de pontos.

Para o cálculo computacional da Transformada de Fourier, foi usada no decorrer desta pesquisa, a Transformada Rápida de Fourier (FFT), definida também em Oppenheim e Schafer

² Para o cálculo tanto do tempo em segundos como os valores da frequência é levado em consideração a taxa de amostragem que será considerada adiante neste capítulo.

(1999) como

$$X[k] = \sum_{n=1}^{N-1} x[n]W_N^{nk}, \quad (2.3)$$

sendo, neste caso, $W_N = e^{-i2\pi/N}$.

A figura 7 mostra o mesmo sinal da figura 6, só que agora no domínio da frequência, mostrando assim o espectro do sinal. Pode ser observado no eixo vertical a magnitude em decibel³ (dB). No eixo horizontal são mostradas as frequências componentes do sinal. Pode-se ainda notar que, para este caso, a componente de maior magnitude, tem frequência de aproximadamente 2.500 hertz, no entanto em todo o espectro há componentes cujas frequências variam de 0 a pouco mais de 4.000 hertz.

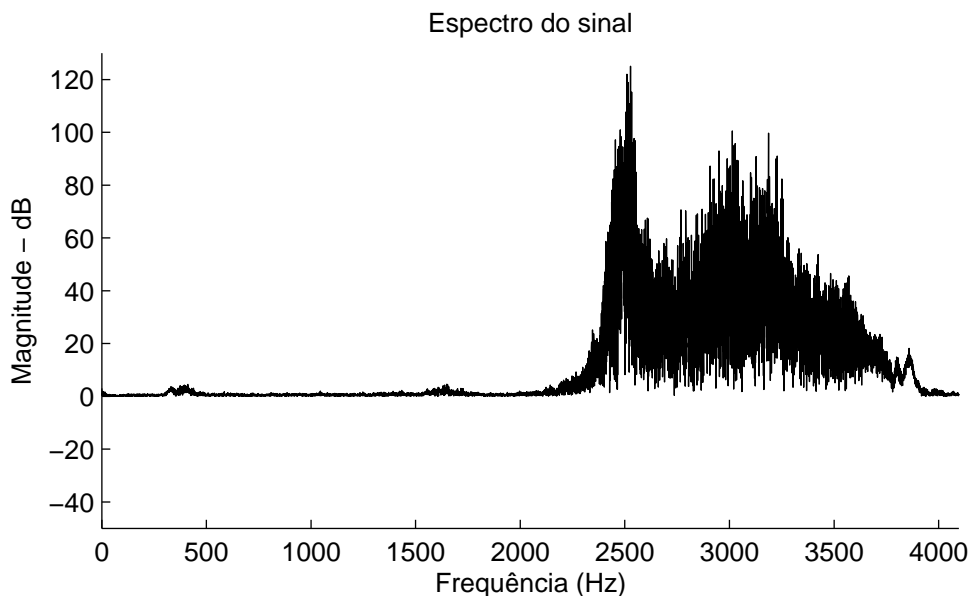


Figura 7 – Exemplo de sinal no domínio da frequência.

2.3 Padronização dos Sinais

Para este trabalho, a fase de coleta do áudio não foi realizada diretamente no *habitat* da ave, mas sim através de buscas em bancos de dados disponíveis na Internet (WIKIAVES, 2011b; XENO, 2011), assim foi possível dar maior atenção aos aspectos diretamente ligados a Processamento Digital de Sinais. Tendo em vista que os sinais analisados são oriundos de várias fontes e formas de gravações diferentes, com taxas de amostragem distintas, formato do arquivo variado e equipamentos também diversos, é necessário padronizar os sinais adquiridos (HAYES, 2006).

³ Decibel é a unidade de medida comumente usada para mensurar a pressão acústica (variedade da intensidade do som) ao longo do tempo (AULETE, 2011).

2.3.1 Formato do Arquivo

Antes de analisar o som em si, é necessário padronizar um formato para o arquivo adquirido. A maioria das vocalizações disponíveis nas bases de dados estão no formato mp3, isto porque é um arquivo considerado pequeno, facilitando assim sua disponibilização na internet (WIKIPEDIA, 2011). A maioria dos ornitólogos, porém, fazem suas gravações originais no formato wav, pois embora este possua um tamanho considerável contempla mais componentes de frequência, deixando o som mais nítido e com maior qualidade (CEO, 2011).

Portanto, ao converter estes arquivos para o formato mp3 podem ocorrer perdas na qualidade do áudio, pois alguns componentes de frequência são descartados para que o arquivo fique menor (AURELIO, 2004). Estes componentes de frequência, uma vez descartados, não podem ser recuperados, mesmo que o arquivo seja novamente convertido para o formato wav (BRANDENBURG, 2009).

Entretanto, como ressaltado em wikipedia (2011), a perda ocasionada por se converter arquivos para o formato mp3 não é tão significativa, chegando a ser imperceptível em alguns casos, isso porque as técnicas utilizadas para comprimir arquivos para o formato mp3 buscam eliminar as partes do som que não são audíveis aos seres humanos. Em adição a isso, para a determinação das características presentes no sinal, com técnicas de DSP, conforme constatado nos resultados, acredita-se que esta pequena variação na qualidade do áudio não é relevante.

O software utilizado para análise dos sons adquiridos e desenvolvimento do classificador foi o MATLAB[®] que, por sua vez, trabalha somente com arquivos no formato wav (MATLAB, 2010). Assim, os arquivos adquiridos em formatos diferentes de wav, tiveram que ser convertidos para este formato. Entretanto, fica ressaltado que a qualidade original proporcionada pelo formato wav não foi restabelecida.

2.3.2 Normalização e Taxa de Amostragem

Para obter uma padronização do sinal, é necessário que o mesmo esteja normalizado. Uma das formas bastante usadas, que também foi aplicada aos sinais testados, é normalizar o sinal em relação ao seu valor máximo de tal forma que ocupem a mesma faixa de amplitude, entre -1 e 1 (SELIN *et al.*, 2007; KWAN *et al.*, 2004). Também o valor médio do sinal pode ser calculado e extraído do sinal para facilitar a análise do espectrograma.

Com relação a taxa, ou frequência, de amostragem, esta se refere a quantidade de amostras retiradas do sinal original por segundo e, de acordo com Smith (1999), a escolha da taxa de

amostragem impõe um limite com relação a quantidade de frequências que o espectro do sinal poderá conter. De acordo com o teorema de Nyquist⁴, os componentes contidos no espectro, terão frequências de, no máximo, metade da frequência de amostragem.

Sendo assim, Hayes (2006), ressalta que no processamento de sinais, podem surgir casos em que seja necessário a alteração da taxa de amostragem, como por exemplo, tratar uma frequência que seja maior do que a metade da que foi coletada. Nestes casos a frequência de amostragem deve ser aumentada. Para isso, é necessário realizar uma operação chamada de interpolação, fazendo com que a escala de tempo seja expandida por um fator L . Entretanto, Hayes (2006) ressalta que a operação de interpolação não recupera a parte perdida do espectro.

De modo geral, como já dito no capítulo introdutório, as aves podem emitir sons numa gama de frequências que podem variar de acordo com o ambiente em que a ave está, podendo até mesmo atingir frequências superiores a 10.000 Hz, chegando seus harmônicos a 18.000 Hz.

Entretanto, como esta pesquisa visa particularmente analisar a vocalização dos tinamídeos, buscou-se, através de testes, determinar qual a faixa de frequência da vocalização da família em geral, constatando que produzem sons entre 300 Hz e 4.000 Hz, aproximadamente. Portanto a taxa de amostragem adotada como padrão foi cerca de 9.000 Hz, garantindo assim uma margem de segurança para a filtragem (SMITH, 1999).

2.4 Filtros Digitais

Uma das operações mais comuns na etapa de pré-processamento é a aplicação de filtros digitais, que podem ser projetados para diversos fins específicos em que seja necessário escolher quais frequências deverão permanecer no sinal e/ou quais devem ser rejeitadas (PAPAMICHALIS, 1987; KIL; SHIN, 1996; PICONE, 1993). Os filtros podem ser projetados para executar quatro operações básicas: passa-baixas, passa-altas, passa-faixa e rejeita-faixa (OPPENHEIM; SCHAFER, 1999).

Segundo Hayes (2006), a seletividade do filtro pode ser verificada pela sua resposta em frequência, obtida pela Transformada de Fourier da resposta ao impulso do filtro. Analisando a figura 8a é possível notar que na resposta em frequência, existe a banda de passagem, ou seja, aquelas frequências que se deseja manter no sinal. Também é possível notar a banda de rejeição, que evidencia a faixa de frequências que foi cortada/atenuada pelo filtro. O intervalo entre a banda de passagem e a banda de rejeição é chamado de banda de transição. Assim, na figura 8 nota-se quatro respostas possíveis para a aplicação de filtros.

⁴ O teorema de Nyquist é explicado em detalhes em Oppenheim e Schaffer (1999, p. 146)

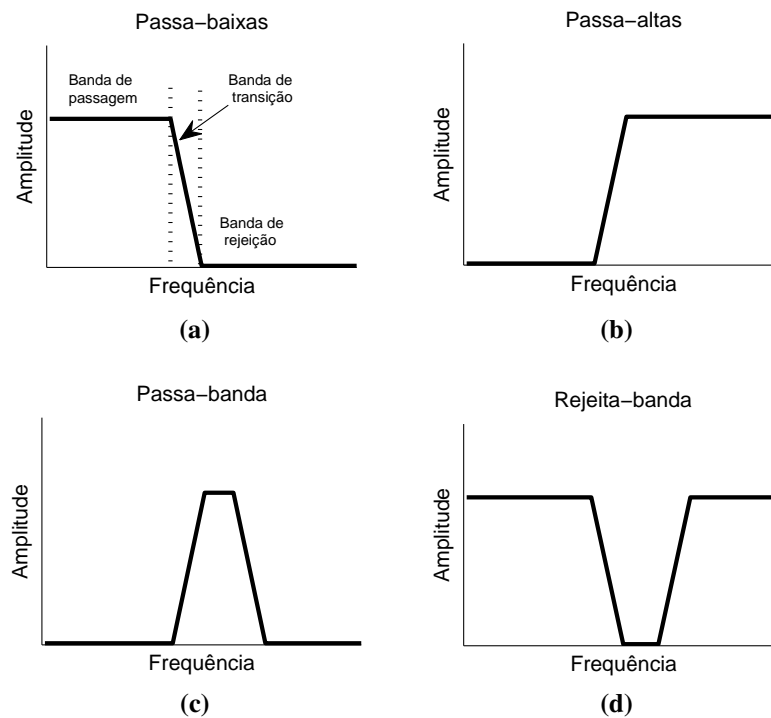


Figura 8 – Resposta da aplicação dos filtros (a) Passa-baixas, (b) Passa-altas, (c) Passa-banda e (d) Rejeita-banda. (Adaptada de Smith (1999, p. 268)).

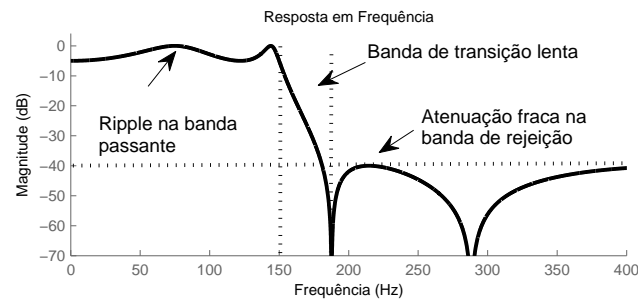
Smith (1999), explicando sobre os quatro tipos de filtros, menciona que filtros passa-baixa (figura 8a) eliminam do sinal qualquer frequência acima de um valor de corte especificado, enquanto que no filtro passa-alta (figura 8b), o contrário acontece, pois é permitido somente a passagem de frequência acima de um ponto de corte. No filtro passa-banda (figura 8c) uma frequência inicial e uma final são escolhidas para que apenas frequências neste intervalo definido passem pelo filtro, enquanto que no filtro rejeita-faixa (figura 8d) um intervalo é definido para ser eliminado do sinal.

De acordo com Haykin e Veen (2001), um filtro ideal seria aquele que deixa passar todos os componentes de frequência que estão na faixa de passagem, não deixando distorção, e por sua vez, eliminam todos os elementos de frequência existentes na faixa de rejeição. Para isso a transição entre a banda de passagem e a banda de rejeição deve ser abrupta.

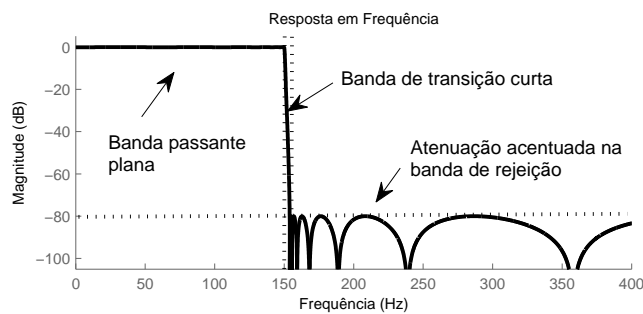
No entanto, Haykin e Veen (2001) bem como Smith (1999), deixam claro que esta implementação ideal é impraticável. Deve-se portanto, tolerar um nível aceitável de distorção e desvios da implementação ideal. Assim, na resposta obtida com a aplicação de um determinado filtro é possível então medir e avaliar o resultado de alguns parâmetros que fornecem dados para decidir qual implementação de filtro é mais indicada a determinada aplicação.

Neste respeito, Smith (1999), classifica alguns parâmetros como ruins e bons, quando uti-

lizado um filtro passa-baixas, conforme pode ser visto na figura 9. Nota-se na figura 9a a existência de *ripple* na banda passante, uma transição lenta e uma baixa atenuação na banda de rejeição – todas características ruins para um filtro. Já na figura 9b, nota-se uma boa implementação, pois há uma banda passante bem plana, transição rápida e atenuação acentuada na banda de rejeição.



(a) Parâmetros ruins



(b) Parâmetros bons

Figura 9 – Parâmetros que avaliam o desempenho de um filtro passa-baixas.

Portanto, é preciso analisar bem a resposta do filtro que se pretende utilizar, para que o resultado seja bem próximo do esperado (HAYKIN; VEEN, 2001). Para isso existem diversas implementações de filtros que podem ser utilizadas tanto para passa-altas, passa-banda, passa-baixas ou rejeita-banda. Serão descritas a seguir algumas destas implementações.

2.4.1 Filtro *Butterworth*

De acordo com Oppenheim e Schaffer (1999), este tipo de filtro é caracterizado por ter uma resposta em frequência bem plana e não possuir *ripples* (ondulações) tanto na banda passante como na banda de rejeição. Os filtros *Butterworth* ainda apresentam um decaimento muito lento na banda de transição, exigindo alta ordem, se for necessário banda de transição mais curta. Entretanto é um filtro que apresenta linearidade e suavidade em comparação com os outros filtros.

Um exemplo de resposta do filtro *Butterworth* é mostrado na figura 10, em que foi implementado um filtro passa-baixas, com frequência de corte de 200 Hz e taxa de amostragem de 1.000 Hz. É possível perceber o decaimento suave deste filtro bem como as diferentes respostas em frequência, dependendo da ordem escolhida para implementação do filtro digital.

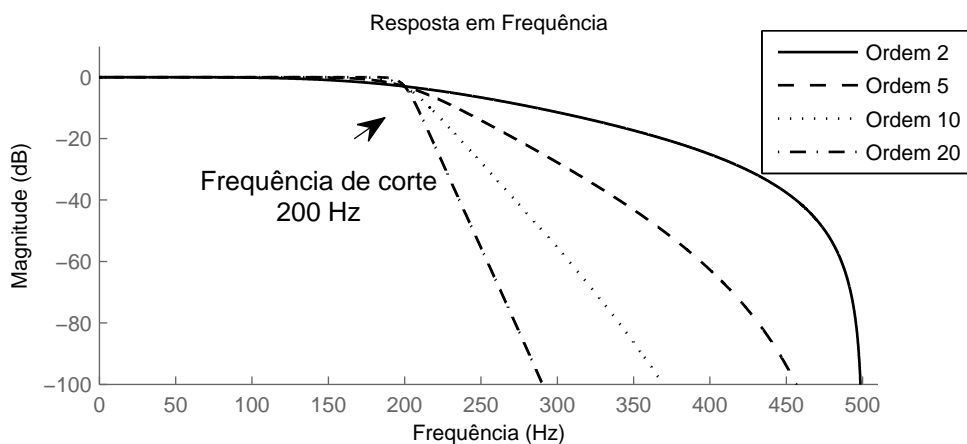


Figura 10 – Resposta do filtro *Butterworth*.

2.4.2 Filtro *Chebyshev*

Os filtros *Chebyshev* apresentam inclinação pouco mais acentuada na banda de transição, ou seja, a passagem entre a banda passante e a banda de rejeição acontece de forma mais rápida, isso sem a necessidade de uma ordem alta. Entretanto, os filtros *Chebyshev* apresentam *ripples* na banda passante, causando distorção dos sinal (SMITH, 1999).

A figura 11 mostra a resposta de um filtro *Chebyshev* para um filtro passa-baixas com frequência de corte de 200 Hz e taxa de amostragem de 1.000 Hz. É notada presença de *ripples* na banda passante, bem como inclinação pouco mais acentuada na banda de transição em relação ao filtro *Butterworth*.

2.4.3 Filtro *Chebyshev* Invertido

Os filtros *Chebyshev* Invertido apresentam ganho plano na banda passante e *ripple* na banda de rejeição. Também possuem inclinação acentuada na banda de transição, exigindo baixa ordem para uma resposta satisfatória (OPPENHEIM; SCHAFER, 1999).

A figura 12 mostra a resposta de um filtro *Chebyshev* Invertido com frequência de corte de 200 Hz, 80 dB para *ripple* na banda de rejeição e taxa de amostragem de 1.000 Hz. Pode-se notar uma banda de passagem bem plana, uma rápida banda de transição, para ordens mais

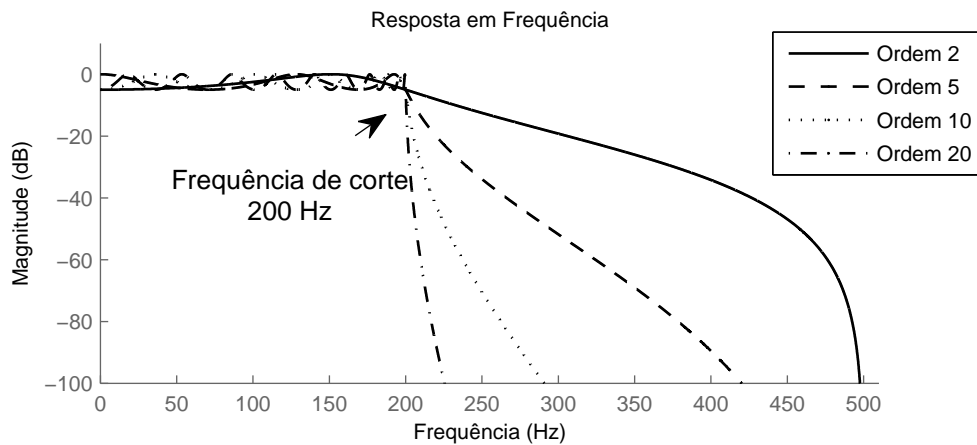


Figura 11 – Resposta do filtro Chebyshev.

elevadas e *ripple* na banda de rejeição. É importante ressaltar que, como o filtro é *Chebyshev* invertido, a inclinação da banda de transição tem início bem antes da frequência de corte, finalizando exatamente no valor estabelecido para o *ripple*. É possível notar isso na figura 12, em que para todas as ordens usadas, a banda de transição finaliza exatamente em -80 dB, entretanto quanto mais alta a ordem, mais próxima da frequência de corte fica o início da banda de transição.

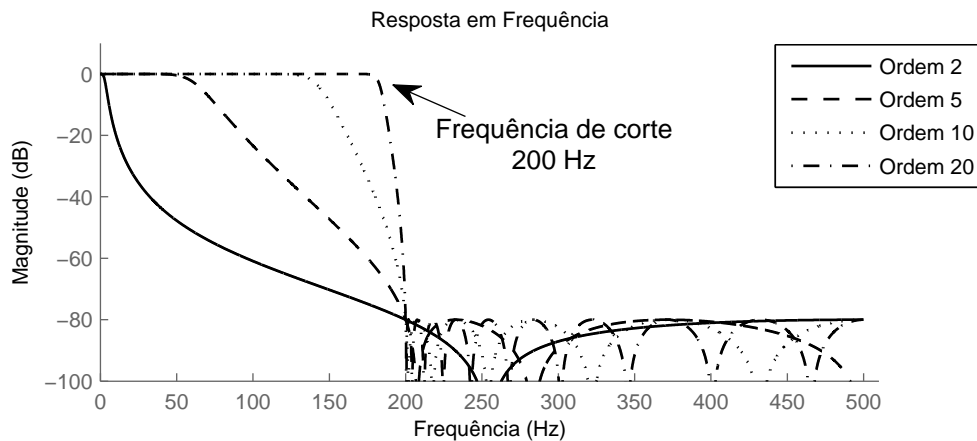


Figura 12 – Resposta do filtro Chebyshev Invertido.

2.4.4 Filtro Elíptico

É possível combinar os filtros *Chebyshev* e *Chebyshev* Invertido de tal forma que coexistam *ripples* tanto na banda de passagem como na banda de rejeição. Quando esta situação ocorre o filtro é denominado elíptico e, para Haykin e Veen (2001) esta situação faz com que a largura da faixa de transição seja menor ainda.

A figura 13 mostra a resposta de um filtro elíptico com frequência de corte de 200 Hz e taxa

de amostragem de 1.000 Hz. Pode-se notar existência de *ripple*, tanto na banda de passagem como na banda de rejeição. Os valores dos *ripples* podem ser parametrizados, neste caso, 5dB na banda passante e 80 dB na banda de rejeição. Os valores destes *ripples* são estabelecidos no projeto do filtro e devem ser bem escolhidos, pois influenciarão na ordem necessária para bom desempenho do filtro. É possível perceber também na figura 13 o quanto a banda de transição é rápida, para a ordem 20, ocorrendo de forma quase que abrupta.

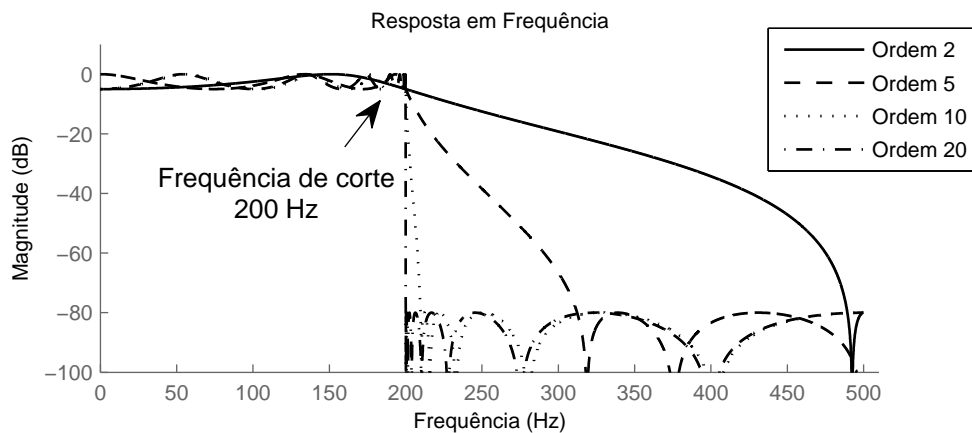


Figura 13 – Resposta do filtro Elíptico.

2.4.5 Escolha do Filtro

Os filtros podem ser parametrizados para que apresentem uma resposta em frequência similar. No entanto, dependendo da implementação, para se atingir a resposta desejada é necessário elevar a ordem do filtro, necessitando de maior custo computacional (SMITH, 1999).

Assim, para escolher o filtro a ser usado no pré-processamento foram feitos testes analisando o sinal e espectro resultante do processo de filtragem e ordem exigida para cada tipo de filtro. A figura 14 mostra o resultado deste teste, em que foi escolhido o sinal da vocalização de uma Azulona (*Tinamus tao*) com muitos distúrbios. Os parâmetros para os filtros foram definidos de tal forma que a resposta em frequência apresentasse, conforme Smith (1999):

- Largura da banda de transição de no máximo 100 Hz;
- *Ripples* de 0.5 dB, na banda passante;
- *Ripples* na banda de rejeição em -80 dB.

Com estas definições, as ordens calculadas para os filtros *Butterworth*, *Chebyshev*, *Chebyshev* Invertido e Elíptico foram, respectivamente, 21, 15, 12 e 9. Nota-se neste caso um melhor de-

sempenho do filtro Elíptico, entretanto o filtro *Chebyshev* Invertido tem vantagem de apresentar banda de passagem plana e a ordem não apresenta grande diferença em relação ao filtro Elíptico.

Conclui-se que todos os filtros conseguiram eliminar os distúrbios e atenuar o ruído fora da faixa de frequência da ave e mantendo as características do sinal. No entanto, o filtro *Chebyshev* Invertido apresentou levemente um melhor desempenho, conseguindo clarear mais o espectro do sinal, conforme análise da figura 14c. Além do mais, este filtro apresenta banda de passagem plana e ordem relativamente baixa. Portanto, para a etapa de pré-processamento foi escolhido como padrão o filtro *Chebyshev* Invertido.

2.5 Interferência de 60 Hz - Filtro Pente

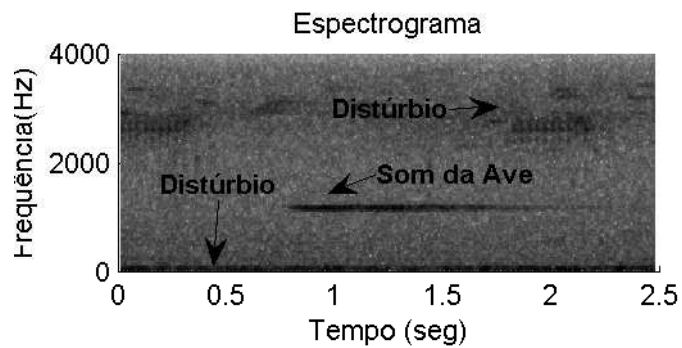
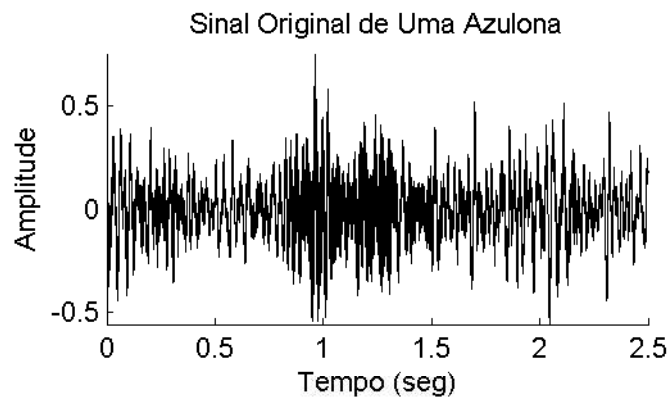
A rede elétrica tem uma frequência de oscilação que, no Brasil, é de 60 Hz. Esta frequência pode interferir nos sinais de áudio gravados se o equipamento utilizado para gravação estiver conectado à rede elétrica ou próximo a redes de alta tensão no momento da gravação (SANCHES, 2003).

Esta interferência causa ruídos, tipo um estalo, iniciando na frequência de 60 Hz se estendendo em grande parte dos seus harmônicos. Este tipo de ruído não pode ser totalmente retirado do sinal original simplesmente com a aplicação dos filtros até então abordados, pois é bem provável que alguns dos harmônicos de 60 Hz coexistirão juntamente com a faixa de passagem desejada (SMITH, 1999).

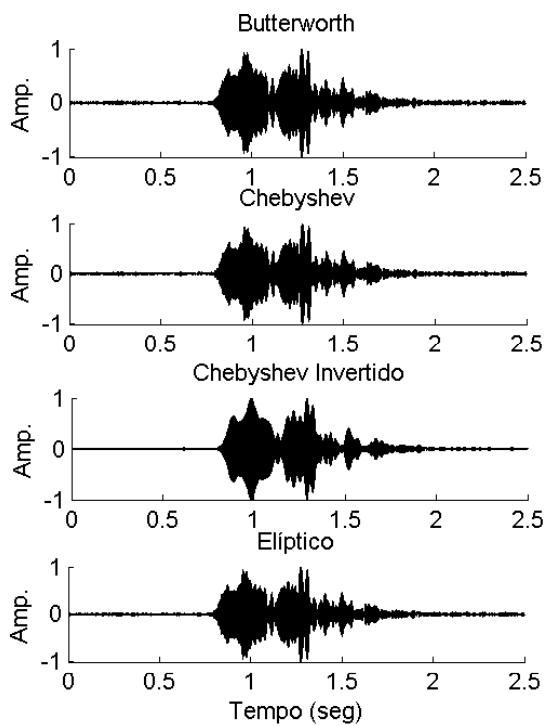
Portanto é necessário utilizar um tipo especial de filtro denominado filtro pente ou *comb filter*. Este tipo de filtro atua como um pente fino, podendo ser projetado para duas finalidades: *peak* - amplifica um conjunto de frequências harmonicamente relacionadas, ou *notch* - atenua um conjunto de frequências harmonicamente relacionadas (MATLAB, 2010).

Para o contexto deste trabalho será utilizado o filtro pente do tipo *notch*, de tal forma que possa atenuar a frequência de 60 Hz bem como de todos os seus harmônicos, deixando o sinal livre deste tipo de interferência.

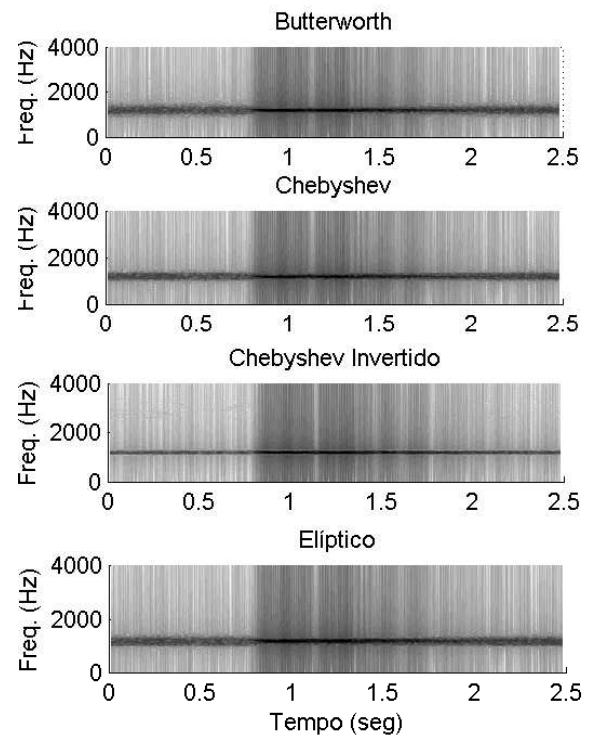
A figura 15 mostra a resposta em frequência de um filtro pente do tipo *notch* para atenuar os componentes de frequência corrompidos pela interferência de 60 Hz. Pode se notar que o filtro inicia seu corte pouco antes de 60 Hz com uma largura que ultrapassa este valor, tendo 60 Hz como seu valor médio. Isso ocorre pelo fato do ruído causado pela interferência não estar exatamente em 60 Hz e seus múltiplos, podendo ocorrer pouco antes ou pouco depois. Esta largura de faixa de corte pode ser definida no projeto do filtro (OPPENHEIM; SCHAFER, 1999).



(a) Sinal e Espectrograma Originais



(b) Sinal Resultante da Filtragem



(c) Espectrograma Resultante da Filtragem

Figura 14 – Resultado do processo de filtragem para a vocalização de um Azulona.

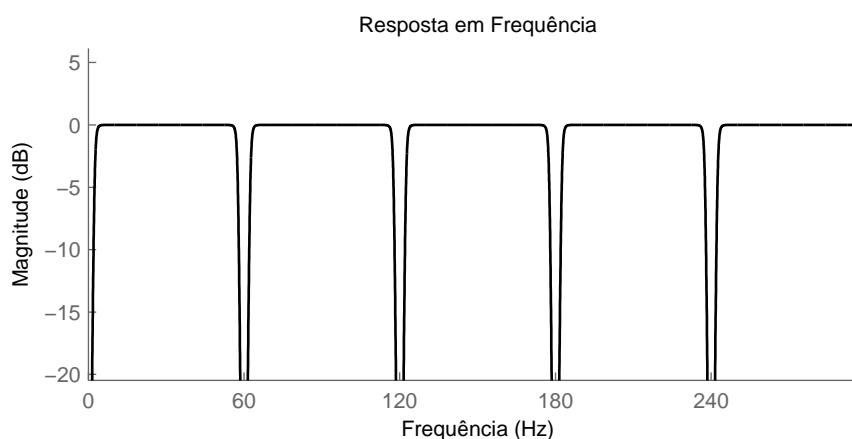


Figura 15 – Resposta do filtro Pente.

2.6 Detecção Automática de Frequências

As detecções das frequências mínima e máxima do sinal devem ocorrer antes da aplicação de um filtro passa banda, pois as frequências mínima e máxima são usadas como parâmetros do filtro, que será aplicado de forma automática. Além disso, os dados de frequência máxima, mínima e mais forte coletados são também armazenados como características da ave, sendo usados posteriormente para o reconhecimento da espécie.

Realizar a estimativa da faixa de frequência visualmente é relativamente simples, bastando visualizar o espectrograma do sinal e verificar o valor aproximado das frequências no instante do som. No entanto, como o objetivo do sistema é funcionar de modo automático, buscou-se uma maneira de encontrar a faixa de frequência através de uma rotina que, automaticamente, analisa os valores do espectrograma do sinal. Os valores do espectrograma são dispostos em uma matriz, de tal forma que é possível verificar a densidade espectral de energia de potência da frequência nas linhas e em relação ao tempo nas colunas.

Ao calcular o espectrograma é preciso definir alguns parâmetros que influenciam na quantidade de pontos existentes na matriz de Tempo-Frequência. Para este trabalho o sinal foi janelado com janelas do tipo *Hamming* de tamanho 256, sobreposição de 128 e o tamanho para o cálculo de FFT de 512, resultando em uma quantidade de pontos suficientes para resultados bem aproximados (KWAN *et al.*, 2004).

Para calcular os valores aproximados da Densidade Espectral de Potência (PSD – Power Spectral Density) é preciso primeiro calcular a Transformada Rápida de Fourier (FFT), para cada segmento de tempo. Então, para cada janela da matriz obtida com cálculo de FFT, é calculada uma estimativa da PSD, que segundo (STOICA; MOSES, 2005), pode ser estimada de

forma simplificada via periodograma, calculando

$$\hat{\phi}_p(w) = \frac{1}{N} \left| \sum_{t=1}^N y(t) e^{-iwt} \right|^2, \quad (2.4)$$

sendo $\hat{\phi}_p(w)$ a densidade de potência obtida, $\sum_{t=1}^N y(t) e^{-iwt}$ o DTFT (Discrete Time Fourier Transform) de $y(t)$, t um determinado ponto, N o número total de pontos.

Após a geração da matriz contendo a PSD para cada segmento, é adotada a seguinte lógica para encontrar as frequências: primeiramente encontra-se o valor máximo da matriz e a linha em que este valor ocorre, em seguida, a partir deste ponto, é feita uma análise de cada linha acima e abaixo, calculando o valor máximo da linha, este valor tende a diminuir cada vez mais. Assim, quando o valor máximo encontrado na linha diminui cerca de 70% do valor máximo da matriz, a busca é encerrada e verificadas as frequências nos instantes de parada. Com esta metodologia é estimada a frequência mais intensa, onde a PSD máxima foi encontrado, a frequência mínima, no instante de parada abaixo do valor máximo e a frequência máxima, no instante de parada acima do valor máximo.

Entretanto, para que esta técnica funcione bem, o som da ave analisada deve conter a maior intensidade presente no sinal, ou seja, nem o ruído, nem outros sons podem ter um volume maior que o da ave em análise.

O valor de limiar adotado foi 70%, escolhido depois de vários testes e análise de resultados, sendo o valor que apresentou um melhor desempenho para a maioria dos sinais analisados. Entretanto, para alguns casos, dependendo do nível de ruído no sinal, este valor pode ser ajustado, para um melhor desempenho.

2.7 Estimativa do Tempo de Canto

Com base no espectrograma resultante da filtragem os valores da PSD Tempo-Frequência são novamente analisados, porém, com objetivo de realizar a segmentação das regiões de som e silêncio. Primeiramente é encontrado o valor da PSD relacionado à frequência mínima detectada e este valor é usado então como limiar. A partir de então é analisado cada segmento de tempo e verificado qual o seu valor máximo, criando um vetor com marcações das regiões de som e silêncio, definido como

$$y[i] = \begin{cases} 1, & \max(P(:, k)) \geq \text{limiar} \\ 0, & \max(P(:, k)) < \text{limiar} \end{cases} \quad (2.5)$$

em que 0 representa silêncio e 1 representa som, P é a matriz de PSD e a notação $P(:, k)$ indica que estão sendo analisadas todas as linhas de uma determinada coluna k da matriz P . Com o vetor y gerado, este é analisado para verificar se as regiões marcadas como silêncio e som estão de acordo com limites pré-definidos. Para este trabalho foi definido um tempo superior a 2 segundos de pausa para ser considerado silêncio e 0.5 segundo de som para ser considerado canto, isto por que foi observado que entre os tinamídeos não há espécie que tenha, em um mesmo canto, pausas de mais de 2 segundos.

A figura 16 demonstra o resultado da técnica para o som de uma Perdiz com duas regiões de som. Percebe-se que neste caso a segmentação funcionou com estimativa bem precisa das regiões de som, formando duas partes distintas de igual tamanho. O tempo de canto é então armazenado, servindo também como característica de reconhecimento.

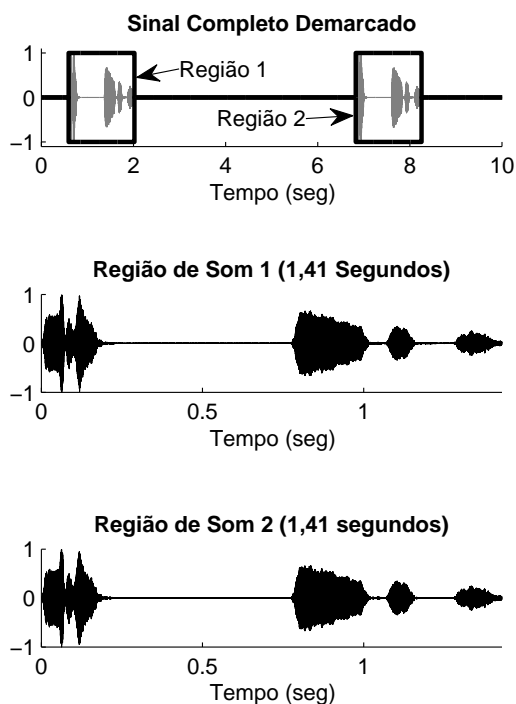


Figura 16 – Exemplo da detecção de regiões de som para o canto de uma Perdiz, sendo detectadas duas regiões de som de igual tamanho.

2.8 Método de Reconhecimento – Vizinho Mais Próximo

Este método estatístico é bastante eficaz ao avaliar as características do sinal de entrada em relação às características previamente incluídas em uma classe de treinamento, calculando

qual a distância do valor de entrada em relação a cada um dos valores armazenados. O sinal de entrada é então classificado baseado no vizinho mais próximo, ou que teve a menor distância calculada (KOHN, 1998).

Para o cálculo desta distância pode ser usado a distância euclidiana, definida em Jelinek (1998, p. 7), como

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.6)$$

Assim, vários indivíduos da mesma espécie são selecionados sendo calculada a média para cada uma das características desejadas, formando uma matriz de treinamento

$$Mt = \begin{bmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1n} \\ \mu_{21} & \mu_{22} & \dots & \mu_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n1} & \mu_{n2} & \dots & \mu_{nn} \end{bmatrix} \quad (2.7)$$

onde cada linha representa uma espécie e cada coluna representa uma característica diferente.

Desta forma, ao realizar o procedimento de reconhecimento de um sinal de entrada desconhecido, suas características são extraídas e calculada a distância euclidiana para cada entrada na matriz de treinamento, sendo então determinada a espécie do sinal de entrada.

Entretanto, Russel e Norvig (2004) apresentam um grande problema ao se usar a distância euclidiana quando cada dimensão do espaço está medindo algo diferente, como é o caso dos valores referentes às frequências das vocalizações dos tinamídeos que podem apresentar valores acima de 3.000 Hz e o tempo de canto abaixo de 1 segundo. Para resolver este problema Russel e Norvig (2004, p. 711) sugerem medir o desvio padrão de cada grupo e expressar os valores de características como múltiplos do desvio padrão para esta característica. Esta métrica se chama distância de Mahalanobis, ou distância euclidiana normalizada, definida como

$$d(x, y) = \sqrt{\sum_{i=1}^n \left(\frac{(x_i - y_i)}{\sigma_i} \right)^2} \quad (2.8)$$

onde σ é o desvio padrão do grupo amostral.

Portanto, para cada entrada da matriz de treinamento deve ser calculado o desvio padrão e gerada uma matriz de desvios padrões como segue

$$M\sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nn} \end{bmatrix} \quad (2.9)$$

Assim, dadas características de entrada desconhecidas, x_1, x_2, \dots, x_n , é calculada a distância euclidiana normalizada, conforme equação 2.8, dos valores de entrada em relação a cada linha da matriz de treinamento, obtendo uma matriz de distâncias M_d em que cada linha representa a distância do sinal de entrada em relação à aquela espécie em particular. Com isso, é encontrado o mínimo valor em M_d e a provável classificação da espécie é determinada.

$$Md = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \quad (2.10)$$

Para Russel e Norvig (2004, p. 712), ao utilizar a distância euclidiana normalizada, é levado em consideração o desvio padrão das características, tornando esta uma eficiente técnica de reconhecimento para um pequeno número de dados. Para Jelinek (1998) e Schurmann (1996) este método estatístico exige muitas amostras já classificadas (treinamento) e ainda é preciso que as características entre as espécies sejam bem distintas para um bom funcionamento.

Portanto, esta metodologia de treinamento e classificação foi escolhida para realizar o reconhecimento de aves da família dos tinamídeos, primeiramente por se tratar de uma família de aves relativamente pequena, além disso a facilidade na aquisição de vocalizações proporcionou ter diversas amostras da mesma espécie, além de que as características extraídas apresentam valores distintos para cada espécie. Todas estas são características que apontam para um bom funcionamento na aplicação da técnica de vizinho mais próximo, calculando a distância euclidiana normalizada pelo desvio padrão.

2.9 Conclusão

Neste capítulo foram apresentados as técnicas usadas no pré-processamento dos sinais e a técnica de reconhecimento a ser utilizada no sistema de reconhecimento automático. Foi indicado que o sinal deve ser padronizado e que deve ser usado o filtro pente somente quando o sinal está corrompido com interferência da rede elétrica. Concluiu-se ainda que o filtro *Chebyshev*

Invertido é adequado para aplicação de um filtro passa-banda. Foi também proposta uma metodologia para a determinação automática da faixa de frequência e regiões de sons, tomando como ponto de partida a análise Tempo-Frequência, calculada através do espectrograma do sinal.

Os parâmetros de frequência e tempo de canto extraídos na fase de pré-processamento serviram como características para o treinamento e reconhecimento, que foram feitos usando a técnica do vizinho mais próximo. No próximo capítulo são mostrados os experimentos e resultados obtidos para cada etapa do sistema de reconhecimento.

Experimentos e Resultados


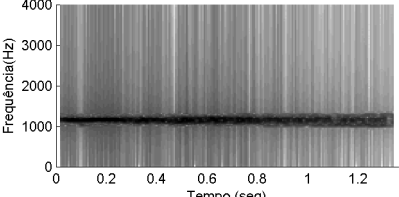

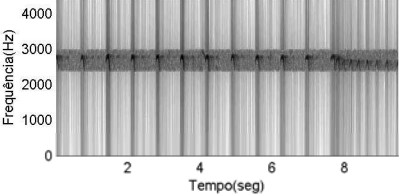

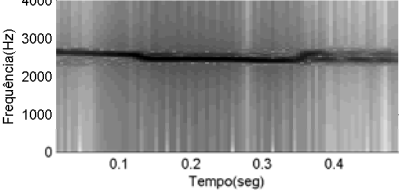

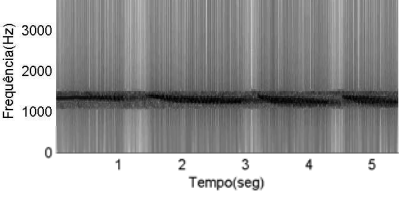

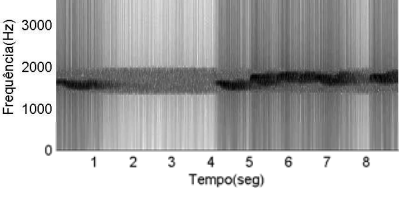

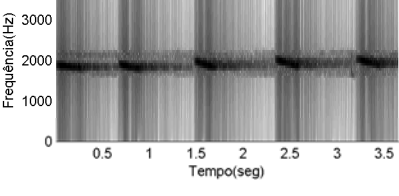
3.1 Introdução

Neste capítulo serão apresentados os resultados do sistema proposto para reconhecimento automático dentro das três fases de um sistema de reconhecimento de padrões: pré-processamento, treinamento e reconhecimento. Para realização dos experimentos do sistema de classificação automática foram usados cerca de 200 sinais referentes a 11 espécies de aves da família Tinamidae, listadas na tabela 2. Estas espécies foram selecionadas por possuírem maior quantidade de vocalizações disponíveis nas bases de dados.

A tabela 2, na primeira coluna, mostra informações de identificação da espécie: abreviatura do nome científico da ave, usada nas demais tabelas, nome científico, nome comum, peso aproximado da espécie em idade adulta, tamanho aproximado da espécie em idade adulta e ambiente(s) que a ave pode ser encontrada. Em seguida são exibidos a imagem e o espectrograma de um canto completo da espécie.


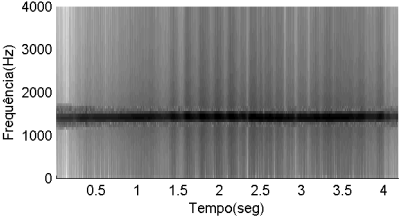

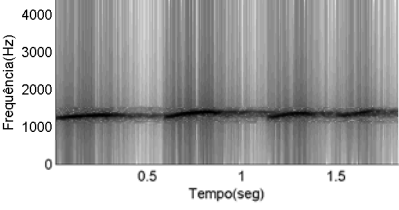

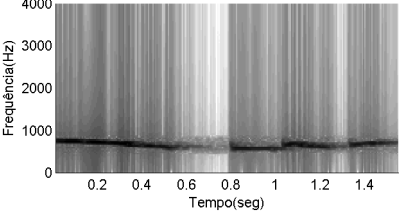

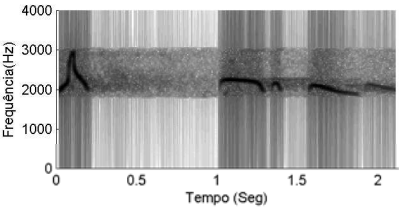

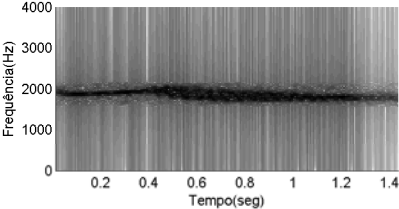
Os resultados apresentados em algumas tabelas usam as seguintes abreviações: FMín (Frequência Mínima, em Hz), FMáx (Frequência Máxima, em Hz), FInt (Frequência mais forte, em Hz), MédC (Média do tempo das regiões de canto, em segundos), μ (Média do grupo), σ (Desvio padrão do grupo), c_v (Coeficiente de Variação), IVis (Inspeção Visual), DAut (Detecção Automática) e Dif (Diferença).

Tabela 2 – Espécies da Família Tinamidae estudadas.

Abreviatura Nome Científico Nome Comum Tamanho/Peso Ambiente(s)	Imagem	Espectrograma
<p><i>TINTAO</i> <i>Tinamus tao</i> Azulona (52cm/1,9Kg) floresta</p>		
<p><i>NOTMAC</i> <i>Nothura maculosa</i> Codorna-amarela (23cm/300g) cerrado/campo</p>		
<p><i>NOTBOR</i> <i>Nothura boraquira</i> Codorna-do-nordeste (27cm/250g) cerrado/campo</p>		
<p><i>TINMAJ</i> <i>Tinamus major</i> Inhambú-açu (41cm/1,05Kg) floresta</p>		
<p><i>CRYVAR</i> <i>Crypturellus variegatus</i> Inhambu-anhangá (28cm/400g) floresta</p>		
<p><i>CRYBAR</i> <i>Crypturellus bartletti</i> Inhambu-anhangá (27cm/400g) floresta</p>		

Continua...

Tabela 2 – continuação...

Abreviatura Nome Científico Nome Comum Tamanho/Peso Ambiente(s)	Imagem	Espectrograma
<p><i>CRYSTR</i> <i>Crypturellus strigulosus</i> Inhambu-relógio (28cm/400g) floresta</p>		
<p><i>CRYUND</i> <i>Crypturellus undulatus</i> Jaó (31cm/800g) floresta</p>		
<p><i>CRYNOC</i> <i>Crypturellus noctivagus</i> Jaó-do-sul (33cm/600g) floresta</p>		
<p><i>RHYRUF</i> <i>Rhynchotus rufescens</i> Perdiz (37cm/800Kg) cerrado/campo</p>		
<p><i>CRYSOU</i> <i>Crypturellus soui</i> Tururim (23cm/230g) floresta</p>		

Fonte dos dados de peso e tamanho: Wikiaves (2011b)

Fonte dos dados sobre ambiente(s): Perlo (2009)

Fonte das imagens: Planet (2012)

3.2 Pré-processamento

A fase de pré-processamento foi implementada tomando por base a figura 5 (seção 2.1). Serão destacadas nas subseções seguintes os resultados do processo de filtragem, extração da faixa de frequência e determinação do tempo de canto.

3.2.1 Detectando e Eliminando Interferência de 60 Hz

A interferência da rede elétrica pode corromper o sinal até mesmo em componentes coexistentes na faixa de frequência da ave. Para detectar se o sinal está corrompido com interferência de 60 Hz é feito um teste verificando qual o componente de frequência presente no sinal com maior magnitude e, se estiver em torno de 60 Hz, é feita a aplicação do filtro pente.

Para ilustrar o procedimento da aplicação do filtro pente e analisar o resultado, foi escolhido um sinal da vocalização de uma Perdiz (*Rhynchotus rufescens*) em que a interferência de 60 Hz é bastante evidente.

A figura 17 mostra o sinal da Perdiz e o seu espectro. É notado que a componente mais evidente no espectro está em torno de 60 Hz, evidenciando assim a interferência da rede elétrica. Pode-se notar ainda que há várias componentes de frequência que são mais evidentes no espectro, estes são os harmônicos de 60 Hz. Este tipo de interferência é tão evidente, com

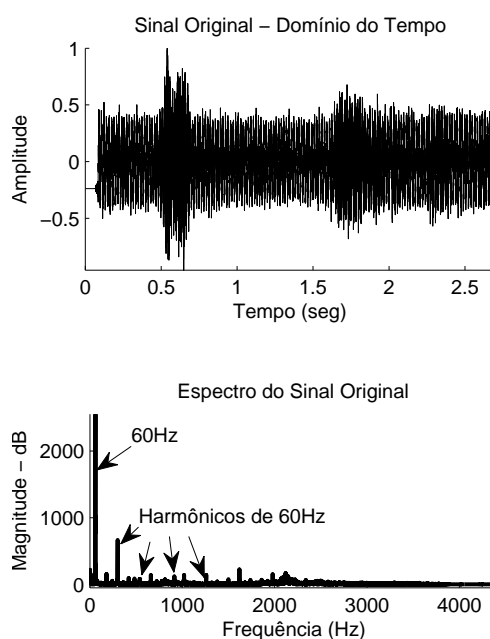


Figura 17 – Sinal e espectro de uma Perdiz, evidenciando a interferência de 60 Hz e seus harmônicos.

magnitudes tão elevadas, que o espectro do canto da ave praticamente não pode ser observado.

A figura 18 mostra o sinal e o espectro da Perdiz após a aplicação do filtro pente. É nítido como o sinal agora está com menos ruído e seu espectro livre dos picos de 60 Hz e dos seus harmônicos.

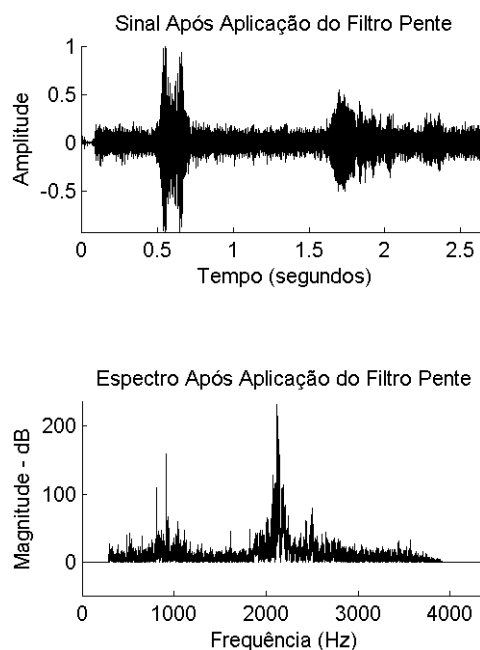


Figura 18 – Sinal e espectro de uma Perdiz, evidenciando a clareza do sinal após a aplicação do filtro pente, que eliminou a interferência de 60 Hz e seus harmônicos.

Portanto, fica claro a importância da aplicação do filtro pente, pois o sinal original da ave é preservado e as frequências indesejadas de 60 Hz e seus harmônicos são bem atenuadas. O sinal estando livre desta interferência, tornam mais exatos e confiáveis os parâmetros detectados automaticamente.

3.2.2 Determinação de Frequências

O procedimento para detecção automática das frequências mais forte, mínima e máxima foi descrito na seção 2.6, indicando a análise Tempo-Frequência para determinação das características.

Para exemplificar o funcionamento deste procedimento será analisada a figura 19. No espectrograma as partes mais escuras da figura evidenciam o canto da ave, sendo a parte mais intensa a frequência mais forte. Nos limites inferiores e superiores é possível constatar as frequências mínima e máxima que são, respectivamente, aproximadamente 1.800 Hz e 2.600 Hz. Ao aplicar a detecção automática foram identificadas como frequência mínima, 1.808 Hz,

máxima, 2.590 Hz e frequência mais forte, 2.136 Hz. Encontrar a frequência mais forte via inspeção visual é relativamente impreciso, pois aparentemente toda a faixa de frequência tem a mesma intensidade, sendo mais prático encontrar esta frequência computacionalmente.

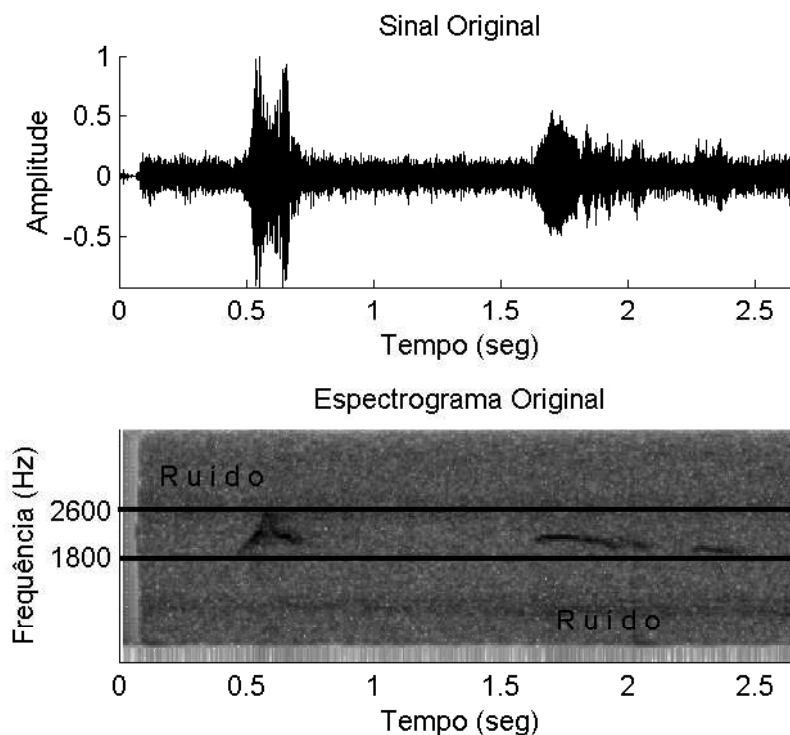


Figura 19 – Sinal e espectro de uma Perdiz, evidenciando a inspeção visual do espectrograma e constatação da faixa de frequência.

Para cada uma das 11 espécies estudadas foram feitas inspeções visuais para identificar as frequências mínima e máxima. Para cada espécie foram calculadas, de forma automática, estas frequências e encontrada a média. A comparação pode ser vista na tabela 3. Para a frequência mais forte considerou-se apenas a detecção automática.

Pode-se concluir da tabela que a detecção automática é satisfatória, mesmo nos casos em que a diferença aparentemente é grande, por alguns fatores: na inspeção visual a verificação é aproximada, sendo que uma pequena variação dá uma grande diferença em Hz. Além disso em ambas as análises deve-se levar em conta que foi calculada média dos valores, devendo considerar uma variação.

3.2.3 Aplicação do Filtro Passa-banda

Com a determinação da faixa de frequência, esta será usada como parâmetro de entrada para o filtro passa-banda, atenuando assim todos os componentes de frequência abaixo e acima do canto da ave.

Tabela 3 – Comparação entre inspeção (IVis) visual e detecção automática (DAut) das frequências mínimas (FMín), mais intensas (FInt) e máximas (FMáx).

Espécie		IVis	DAut	Dif (%)
TINTAO	FMín	1000	1082,62	8,26
	FInt	--	1196,86	--
	FMáx	1400	1305,91	7,20
NOTMAC	FMín	2300	2472,97	7,52
	FInt	--	2714,58	--
	FMáx	3200	2882,58	11,01
NOTBOR	FMín	2100	2335,16	11,19
	FInt	--	2488,96	--
	FMáx	2800	2678,73	4,52
TINMAJ	FMín	1100	1204,63	12,44
	FInt	--	1348,60	--
	FMáx	1500	1491,33	6,23
CRYVAR	FMín	1300	1375,97	5,07
	FInt	--	1558,35	--
	FMáx	2000	1811,81	10,38
CRYBAR	FMín	1600	1710,45	6,90
	FInt	--	1866,46	--
	FMáx	2200	2041,71	7,75
CRYSTR	FMín	1150	1278,05	9,96
	FInt	--	1393,55	--
	FMáx	1600	1510,79	5,90
CRYUND	FMín	950	1080,30	13,71
	FInt	--	1241,08	--
	FMáx	1550	1413,34	9,66
CRYNOC	FMín	450	560,82	2,40
	FInt	--	740,74	--
	FMáx	800	740,74	8,00
RHYRUF	FMín	1800	1877,02	4,27
	FInt	--	2140,64	--
	FMáx	2600	2468,12	5,34
CRYSOU	FMín	1600	1753,19	9,57
	FInt	--	2004,83	--
	FMáx	2400	2286,83	4,94

A seção 2.4.5 fez uma comparação entre os principais filtros e chegou a conclusão que seria melhor a utilização do filtro *Chebyshev* Invertido. Para corroborar com esta conclusão é mostrado na figura 20 o resultado do processo de filtragem para o canto de uma Perdiz (mesmo sinal da figura 19), já com os parâmetros detectados de forma automática. É possível perceber inicialmente que o sinal está com menos ruído. No espectrograma nota-se que, fora da faixa

de frequência do canto da ave, o ruído foi bem atenuado. Percebe-se também que a faixa de frequência não atenuada contempla, de forma bem aproximada, a faixa de frequência fundamental do canto da ave.

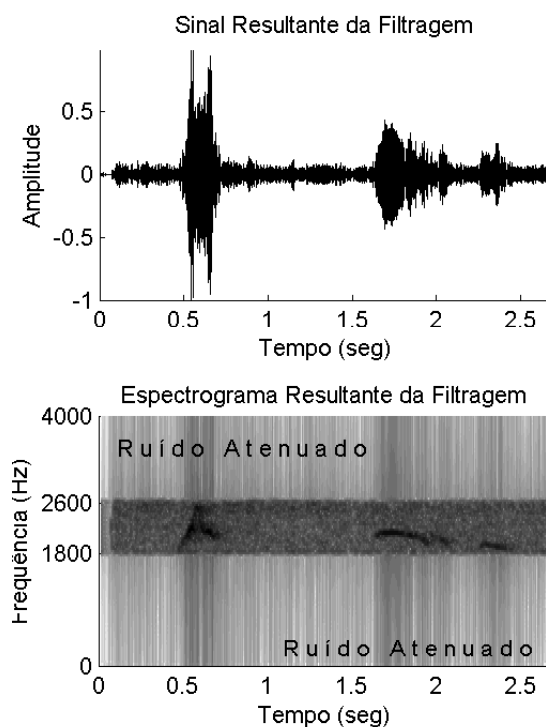


Figura 20 – Sinal e espectro de uma Perdiz, filtrado com parâmetros determinados automaticamente.

3.2.4 Determinação do Tempo de Canto

Uma característica também importante para reconhecimento é o tempo de duração do canto da ave. Esta característica foi também determinada de forma automática, conforme descrito na seção 2.7. Mais uma vez é possível fazer inspeção visual e encontrar o tempo de duração do canto da ave, conforme figura 21, para canto de um Jaó (*Crypturellus undulatus*). Nota-se que o canto inicia em aproximadamente 1 segundo e termina em aproximadamente 3 segundos durando, portanto, cerca de 2 segundos.

A figura 22 mostra o mesmo sinal segmentado de forma automática. Foi detectado que o canto se inicia em 1 segundo e termina em 2,8 segundos, durando cerca de 1,8 segundos, conforme mostra a parte inferior da figura, isolando e ampliando apenas a região de som.

A tabela 4 mostra a comparação de todas as espécies estudadas. É possível constatar que a detecção automática para o tempo de canto, também não apresenta grande diferença em relação à análise visual. Notou-se que, mesmo na análise visual, o tempo de canto apresenta uma

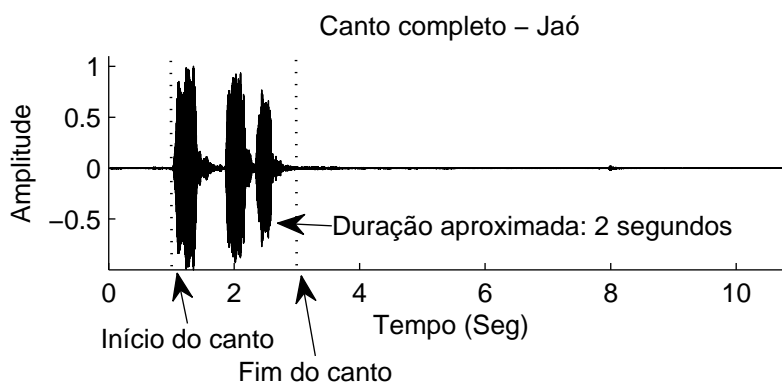


Figura 21 – Inspeção visual para detecção da região de som do sinal de um Jaó.

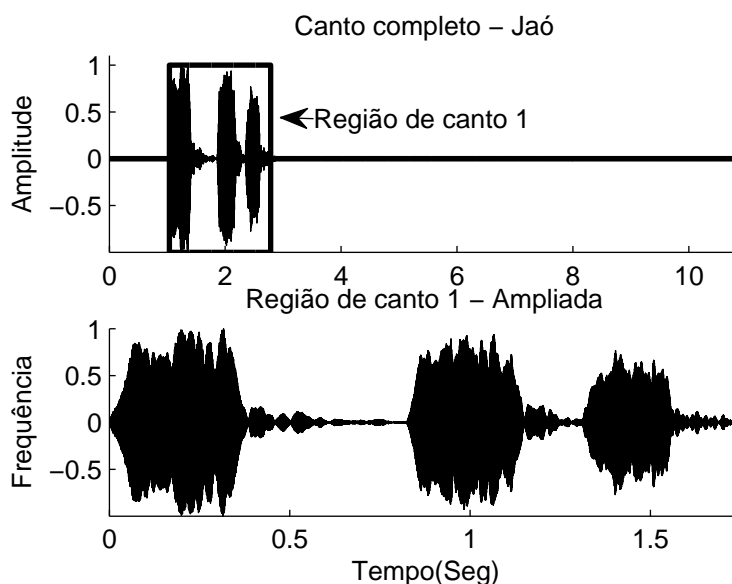


Figura 22 – Detecção automática para encontrar a região de som do sinal de um Jaó.

variação mais acentuada, em alguns casos a mesma ave apresenta tempos de canto diferentes. Entretanto, o tempo de canto ainda é uma característica relevante para o reconhecimento, tendo em vista que os valores são bem distintos entre as diferentes espécies.

3.3 Treinamento

Conforme definido no capítulo 2, será utilizado para treinamento e classificação o método de menor distância, também conhecido como vizinho mais próximo.

A etapa de treinamento para esta metodologia em particular, envolve selecionar alguns sinais que servirão para determinação de valores padrão. Assim, ao realizar o procedimento de

Tabela 4 – Comparação entre inspeção visual (IVis) e detecção automática (DAut) do tempo de canto (MédC).

Espécie		IVis	DAut	Dif (%)
TINTAO	MédC	1,10	0,97	11,81
NOTMAC	MédC	10,00	8,23	17,70
NOTBOR	MédC	0,70	0,53	24,28
TINMAJ	MédC	5,00	3,94	21,20
CRYVAR	MédC	8,10	7,25	10,49
CRYBAR	MédC	5,10	4,20	18,80
CRYSTR	MédC	4,30	4,55	5,81
CRYUND	MédC	2,00	1,93	3,50
CRYNOC	MédC	1,40	1,43	2,14
RHYRUF	MédC	1,70	1,83	7,64
CRYSOU	MédC	1,30	1,53	17,69

classificação, o sinal de entrada é comparado com a tabela de treinamento para que possa ser determinada a espécie da ave. Dos sinais adquiridos, foram escolhidos, aleatoriamente, cerca de 60 para compor a tabela de treinamento.

Inicialmente, alguns sinais de vocalização de uma mesma espécie são agrupados para formar os valores de treinamento da referida espécie, calculando a média, o desvio padrão e o coeficiente de variação. Este procedimento é mostrado nas tabelas 5, 6 e 7.

Tabela 5 – Dados coletados de *Crypturellus undulatus* (Jaó).

Ave	FMín	FInt	FMax	MédC
Jaó1	1085,27	1274,77	1429,80	1,51
Jaó2	1085,27	1240,31	1395,35	2,07
Jaó3	1033,59	1205,86	1378,13	1,52
Jaó4	1068,05	1171,41	1412,58	1,61
Jaó5	1085,27	1223,09	1395,35	1,61
Jaó6	1050,82	1240,31	1429,8	1,80
Jaó7	1136,95	1291,99	1447,03	2,00
μ	1077,90	1235,40	1412,60	1,73
σ	32,77	40,66	24,36	0,23
c_v	3,04	3,29	1,72	13,29

Os valores das médias, desvios padrões e coeficientes de variação para todas as espécies estudadas estão dispostos na tabela 8 em que é possível perceber uma variação baixa, evidenciando assim que a determinação automática das características têm valores aceitáveis e que os

Tabela 6 – Dados coletados de *Crypturellus strigulosus* (Inhambu relógio).

Ave	FMín	FInt	FMáx	MédC
Inhambu1	1218,75	1406,25	1562,5	3,99
Inhambu2	1378,13	1481,48	1567,62	4,75
Inhambu3	1223,09	1343,67	1429,80	4,11
Inhambu4	1343,67	1429,80	1533,16	4,82
Inhambu5	1188,63	1291,99	1378,13	5,23
μ	1270,45	1390,64	1494,24	4,48
σ	84,50	74,10	85,43	0,51
c_v	6,65	5,33	5,72	11,38

Tabela 7 – Dados coletados de *Rhynchotus rufescens* (Perdiz).

Ave	FMín	FInt	FMáx	MédC
Perdiz1	1826,02	2153,32	2411,72	1,64
Perdiz2	1808,79	1946,60	2291,13	1,80
Perdiz3	1808,79	2222,23	2980,20	1,89
Perdiz4	1808,79	1981,05	2170,55	1,91
Perdiz5	1963,83	2205,00	2325,59	1,89
Perdiz6	1998,28	2136,09	2291,13	1,64
Perdiz7	2291,13	2497,85	2670,12	1,74
Perdiz8	2291,13	2497,85	2670,12	1,74
Perdiz9	1912,15	2205,00	2618,44	1,55
μ	1967,66	2205,00	2492,11	1,76
σ	196,56	192,41	258,83	0,13
c_v	9,99	8,73	10,39	7,65

mesmo podem ser usados como valores do treinamento.

É importante ressaltar que os valores base para originar a tabela 8 foram detectados automaticamente e apresentam-se bem homogêneos dentro de cada espécie e distintos em relação a outra espécie diferente. A figura 23 mostra a disposição gráfica por espécie dos parâmetros coletados. Este formato de gráfico foi idealizado exclusivamente para os parâmetros coletados, sendo que no eixo x, as bases representam as frequências mínima e máxima, o pico representa a frequência mais intensa. No eixo y, onde ocorrem os valores de pico, representa o tempo de canto. É possível perceber que, na maior parte dos casos, os valores caracterizam individualmente cada espécie, pois quando alguns dados de frequência apresentam-se similares, diferenciam-se no tempo de canto e vice-versa.

Tabela 8 – Tabela de treinamento para as espécies estudadas.

Espécie		FMín	FMáx	FInt	Médc
<i>TINTAO</i>	μ	1093,90	1326,40	1218,80	0,96
	σ	49,58	59,67	54,28	0,27
	c_v	4,53	4,45	4,50	28,13
<i>NOTMAC</i>	μ	2523,70	2894,10	2764,90	8,79
	σ	255,80	48,72	85,26	0,16
	c_v	10,14	3,08	1,68	1,82
<i>NOTBOR</i>	μ	2242,30	2614,10	2417,50	0,56
	σ	125,51	67,68	60,49	0,17
	c_v	5,60	2,50	2,59	30,36
<i>TINMAJ</i>	μ	1201,60	1490,10	1335,10	4,77
	σ	61,91	35,86	53,56	1,77
	c_v	5,15	4,01	2,41	37,11
<i>CRYVAR</i>	μ	1408,30	1831,80	1663,80	8,03
	σ	35,25	55,37	137,30	0,77
	c_v	2,50	8,25	3,02	9,59
<i>CRYBAR</i>	μ	1632,70	2039,50	1831,90	5,00
	σ	18,46	101,15	69,66	1,53
	c_v	1,13	3,8	4,96	30,60
<i>CRYSTR</i>	μ	1270,50	1494,20	1390,60	4,58
	σ	84,50	85,43	74,10	0,51
	c_v	6,65	5,33	5,72	11,14
<i>CRYUND</i>	μ	1077,90	1412,60	1235,40	1,73
	σ	32,77	24,36	40,66	0,23
	c_v	3,04	3,29	1,72	13,29
<i>CRYNOC</i>	μ	536,89	854,15	716,34	1,42
	σ	63,48	36,62	54,53	0,15
	c_v	11,82	7,61	4,29	10,56
<i>RHYRUF</i>	μ	1967,70	2492,10	2205,00	1,75
	σ	196,56	258,83	192,41	0,12
	c_v	9,99	8,73	10,39	6,86
<i>CRYSOU</i>	μ	1864,80	2304,10	2118,90	1,44
	σ	159,37	122,92	151,49	0,27
	c_v	8,55	7,15	5,33	18,75

3.4 Reconhecimento

O procedimento para classificação da espécie é ilustrado na figura 24. Ao receber um sinal de entrada desconhecido, o mesmo passará pelo pré-processamento normalmente e suas características de frequências e tempo de canto são então determinadas e comparadas com cada um dos conjuntos de valores dentro da tabela de treinamento, os valores que mais se aproximarem

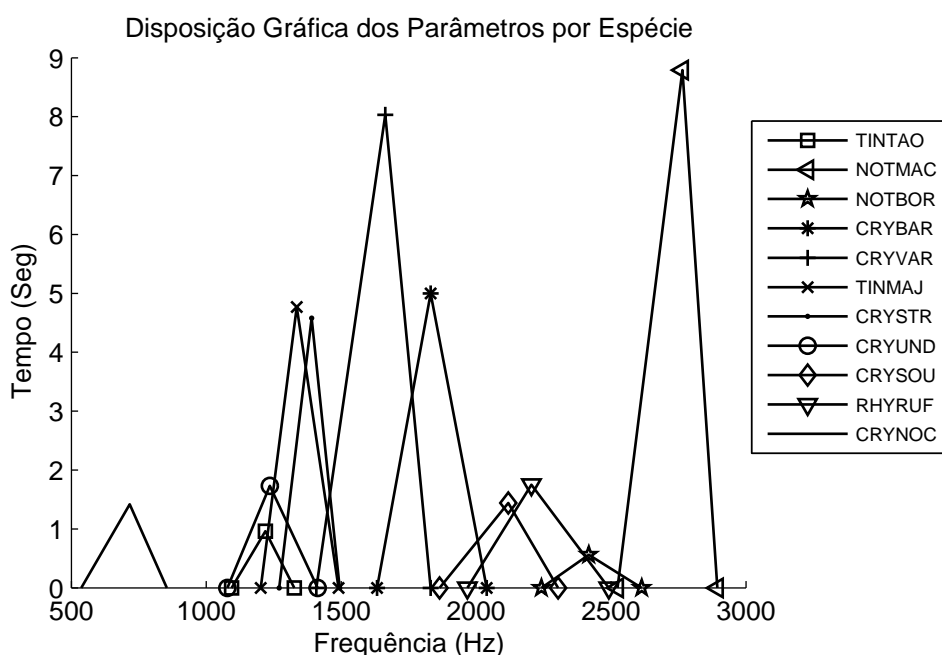


Figura 23 – Disposição gráfica dos parâmetros coletados por espécie - Tempo x Frequência.

aos determinados do sinal de entrada são considerados da espécie provável.

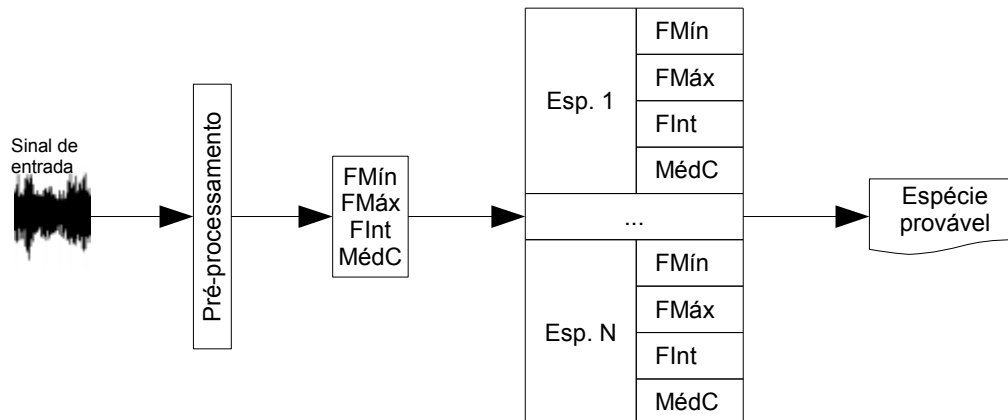


Figura 24 – Sequência do processo de reconhecimento de um sinal de entrada.

Para exemplificar o procedimento, a figura 25 mostra o sinal e espectrograma de um Inhambú-relógio (*Crypturellus strigulosus*), desconhecido para o sistema. Ao determinar as frequências mínima, mais intensa e máxima, obtém-se respectivamente 1.218 Hz, 1.406 Hz e 1.562 Hz. O sinal apresenta duas regiões de canto sendo a primeira determinada com duração de 4,03 segundos e a segunda com 3,95 segundos, tendo 3,99 segundos de média.

Para verificar a aproximação dos valores de entrada com os valores da tabela de treinamento, foi usada a distância Euclidiana, sendo os valores normalizados pelo desvio padrão de cada espécie na tabela de treinamento. Assim, para o Inhambú-relógio em processo de reconhe-

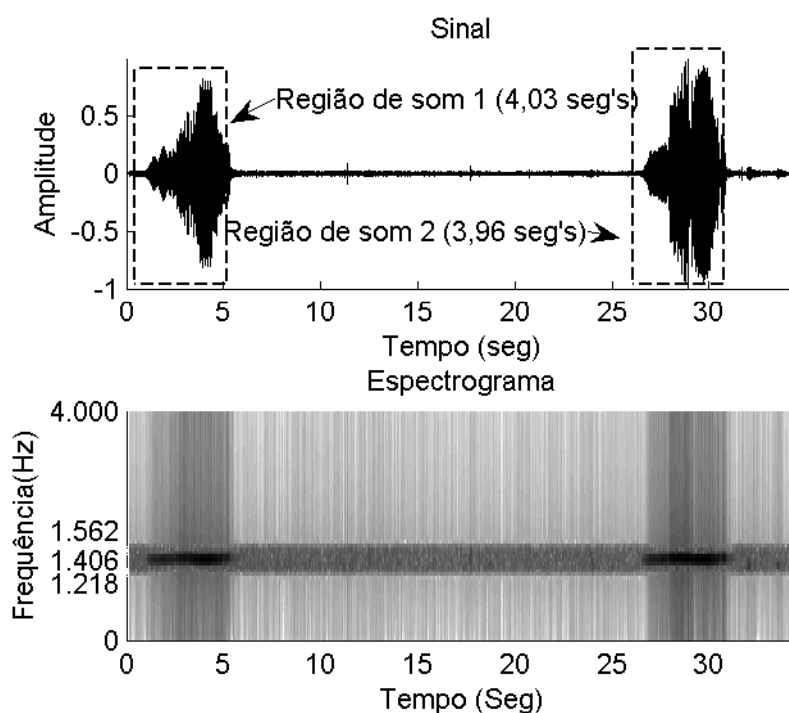


Figura 25 – Determinação das características de um sinal de entrada.

cimento, conforme figura 25, obtém-se a tabela 9 de distâncias. Percebe-se que o menor valor calculado para a tabela 9 é 1,53, sendo esta linha correspondente à entrada do Inhambú-relógio na tabela de treinamento. Neste caso a abordagem de menor distância obteve sucesso.

Tabela 9 – Distâncias calculadas para o exemplo de um *Crypturellus strigulosus*.

Espécie	Distância
<i>TINTAO</i>	12,32
<i>NOTMAC</i>	42,91
<i>NOTBOR</i>	31,18
<i>TINMAJ</i>	2,47
<i>CRYVAR</i>	9,10
<i>CRYBAR</i>	23,71
<i>CRYSTR</i>	1,53
<i>CRYUND</i>	13,01
<i>CRYNOC</i>	30,17
<i>RHYRUF</i>	18,59
<i>CRYSOU</i>	12,78

Graficamente também é possível constatar o resultado, conforme figura 26. Nota-se que os valores encontrados, quando dispostos na gráfico em que se encontram todas as espécies, se aproxima mais dos dados referentes ao Inhambú-relógio.

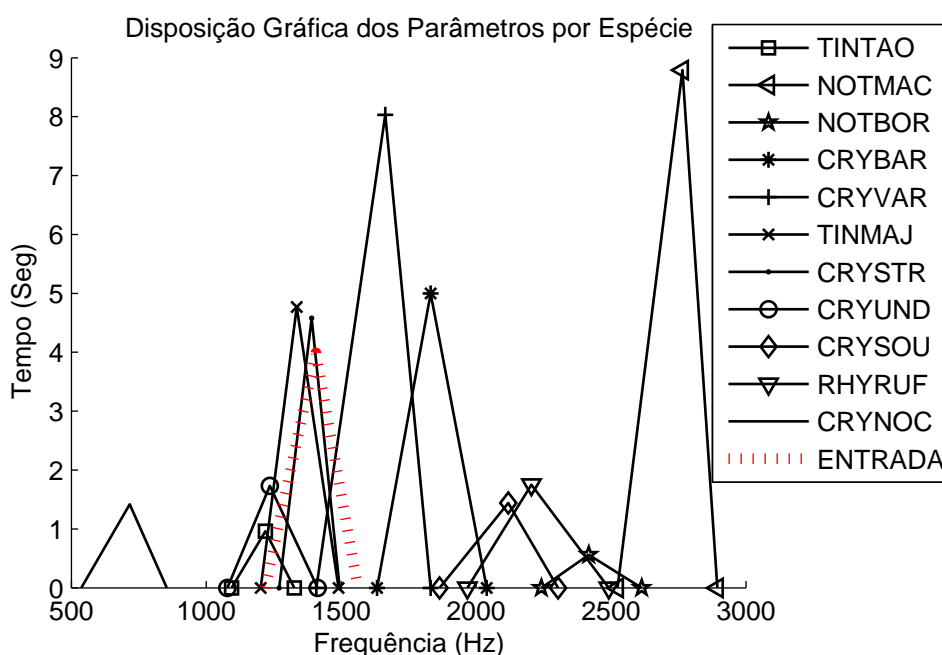


Figura 26 – Disposição gráfica com sinal de entrada.

Tabela 10 – Resultados obtidos na classificação de algumas espécies.

Espécie	Quantidade Testada	Acerto
TINTAO	18	100,00%
NOTMAC	7	100,00%
NOTBOR	6	100,00%
TINMAJ	10	80,00%
CRYVAR	5	80,00%
CRYBAR	5	60,00%
CRYSTR	12	100,00%
CRYUND	30	96,67%
CRYNOC	6	100,00%
RHYRUF	31	96,77%
CRYSOU	6	83,33%
Total	136	94,12%

Os resultados obtidos com esta técnica de reconhecimento podem ser vistos nas tabelas 10 e 11. Nota-se por estas tabelas de resultados que, de modo geral, houve 94,12% de acerto, o que pode ser considerado bastante satisfatório. Entretanto, para algumas espécies, obteve-se um índice de acerto baixo, como no caso de *Crypturellus bartletti*, em que o índice de acerto foi de apenas 60%. Um dos fatores deste baixo índice de acerto pode ser porque o teste foi realizado com poucas espécies, tendo em vista que são sons mais difíceis de se encontrar. Entretanto nota-se na tabela 8, que algumas características são relativamente próximas umas das outras

e, como os desvios padrões são levados em conta, pode ser que uma espécie se confunda com outra. Ainda assim, de modo geral os resultados da técnica são positivos.

Tabela 11 – Matriz de confusão dos sinais testados.

Espécie	<i>TINTAO</i>	<i>NOTMAC</i>	<i>NOTBOR</i>	<i>TINMAJ</i>	<i>CRYVAR</i>	<i>CRYBAR</i>	<i>CRYSTR</i>	<i>CRYUND</i>	<i>CRYNOC</i>	<i>RHYRUF</i>	<i>CRYSOU</i>
<i>TINTAO</i>	18										
<i>NOTMAC</i>		7									
<i>NOTBOR</i>			6								
<i>TINMAJ</i>				8			2				
<i>CRYVAR</i>					4		1				
<i>CRYBAR</i>						3	1			1	
<i>CRYSTR</i>							12				
<i>CRYUND</i>	1							29			
<i>CRYNOC</i>									6		
<i>RHYRUF</i>										30	1
<i>CRYSOU</i>						1					5

Na tabela 11 nota-se mais detalhadamente o resultado para cada tentativa de classificação, evidenciando assim os falsos positivos e falsos negativos. É interessante notar que para os casos em que não houve acerto, a espécie classificada equivocadamente realmente têm parâmetros similares com a espécie correta.

3.5 Conclusão

Neste capítulo foram mostrados os resultados da aplicação dos procedimentos de pré-processamento, treinamento e reconhecimento. Características de frequência mínima, máxima e mais intensa bem como o tempo de canto foram armazenadas em uma tabela de treinamento para que um sinal de entrada pudesse ser testado e classificado. Para calcular a proximidade dos valores de entrada com os valores armazenados no treinamento foi usada a Distância Euclidiana normalizada pelo desvio padrão, o que mostrou ser uma técnica eficiente com acurácia de 94,12%.

Conclusão e Trabalhos Futuros

Esta pesquisa propõe um sistema metodológico para reconhecimento e classificação de aves através da sua vocalização, focando exclusivamente no reconhecimento de aves da família Tinamidae.

O sistema proposto no pré-processamento definiu uma técnica para determinação automática de características de frequência mínima, máxima e mais intensa. Além disso, no pré-processamento foi também feita a segmentação do canto em períodos de som e silêncio.

A técnica de análise tempo-frequência, para determinar as características de frequência e tempo de canto, se mostrou eficaz, pois conseguiu delimitar bem, tanto a faixa de frequência da vocalização da ave, como também o tempo de canto.

Entretanto, tendo em vista que a técnica usa um valor fixo como limiar, 70% da frequência mais intensa, pode não funcionar para espécies de outras famílias. Além disso se houver algum som, em qualquer ponto do sinal, que seja mais intenso que a frequência da vocalização da ave, esta técnica também falha. A metodologia proposta também encontra dificuldades quando o sinal é bastante ruidoso, pois o limiar pode nunca ser atingido.

Mesmo com as dificuldades encontradas, os parâmetros extraídos do sinal foram, na grande maioria das vezes, precisos em relação à inspeção visual, se mostrando bem homogêneos entre si e bem distintos em relação ao sinal de outra espécie.

Assim, para esta etapa de pré-processamento, fica como contribuição uma forma de analisar o espectrograma do sinal como alternativa na busca automática de características da vocalização. Fica como sugestão para trabalhos futuros, buscar meios mais dinâmicos de determinação do limiar, certificação de que a frequência realmente é do canto da ave e também conseguir determinar as características mesmo com presença de forte ruído ou com outras aves da mesma espécie.

As etapas de treinamento e classificação foram feitas usando a técnica do vizinho mais próximo, calculando a distância Euclidiana normalizada pelo desvio padrão. Para esta família de aves, esta técnica foi bastante eficaz, conseguindo 94,12% de acurácia.

Ao utilizar a distância Euclidiana, fica realmente provado que é uma técnica eficiente quando o número de espécies a ser reconhecido é pequeno e os parâmetros utilizados, distintos. Entretanto para um reconhecimento generalizado, esta técnica talvez não funcionasse tão bem.

Portanto fica como sugestão para trabalhos futuros utilizar outras técnicas de reconhecimento, como por exemplo as Redes Neurais Artificiais (RNA), pois esta pode ser usada num âmbito mais generalizado. Uma outra sugestão seria a determinação de mais características do sinal de entrada e a otimização destas características, analisando qual melhor se encaixa em determinada técnica de reconhecimento.

Como maior contribuição espera-se que este trabalho incentive diversas outras pesquisas relacionadas ao reconhecimento de aves e que, de alguma forma, possa auxiliar na monitoração e preservação de aves, pois o número de espécies que entram em ameaça de extinção, aumenta a cada ano.

Referências

- ALEGRINI, M. F. Dissertação de mestrado. *Avifauna como possível Indicador ecológico para os estádios de regeneração da mata atlântica*. São Paulo: USP, 1997.
- ALMEIDA, A. F. Tese de Doutorado. *Avifauna de uma área desflorestada em Anhembi, Estado de São Paulo, Brasil*. São Paulo: USP, 1981.
- ALMEIDA Álvaro Fernando de; ALMEIDA, A. de. Monitoring fauna and habitats in forested areas. *Série Técnica IPEF*, 1998.
- ANTAS, P. de T. Z.; CAVALCANTI, R. B. *Aves Comuns do Planalto Central*. Brasília: Universidade de Brasília, 1988.
- AULETE. *Aulete Digital, Dicionário Contemporâneo da Língua Portuguesa - Software*. CaldasAulete, 2011. Disponível em: <<http://www.auletedigital.com.br/>>. Acesso em: 01/11/2011.
- AURELIO, D. *Desvendando o MP3*. São Paulo: Digerati Books, 2004. ISBN 85-89535-35-5.
- AVIBASE. *Avibase -The World Birds Database*. Avibase, 2011. Disponível em: <<http://avibase.bsc-eoc.org>>. Acesso em: 28/02/2011.
- BIRDLIFE. *Birdlife International - Working together for birds and people*. Birdlife International, 2011. Disponível em: <<http://www.birdlife.org/index.html>>. Acesso em: 01/11/2011.
- BRANDENBURG, K. Mp3 and aac explained. *AES 17th International Conference on High Quality Audio Coding*, 2009.
- BRANDES, T. S. Automated sound recording and analysis techniques for bird surveys and conservation. *Bird Conservation International*, p. S163–S173, 2008.
- BROWN, J. H.; LOMOLINO, M. V. *Biogeografia 2 ed.* [S.l.]: Funpec, 2006.
- CBRO. *CBRO, Comitê Brasileiro de Registros Ornitológicos - Lista de Aves 2011*. Comitê Brasileiro de Registros Ornitológicos, 2010. Disponível em: <<http://www.cbro.org.br>>. Acesso em: 27/02/2010.
- CEO. *CEO - Gravando vozes de aves*. CEO, 2011. Disponível em: <<http://www.ceo.org.br/observ/bioac.htm>>. Acesso em: 03/11/2011.

- CORREIO. Bioacústica identifica novas aves na amazônia. *Artigo Eletrônico Correio do Brasil*, 2010. Disponível em: <<http://correiodobrasil.com.br/bioacustica-identifica-novas-aves-na-amazonia/152369/>>. Acesso em: 28/02/2011.
- DESCOURTILZ, J. T. *História Natural das Aves do Brasil - Notáveis por Sua Plumagem, Canto e Hábitos*. Belo Horizonte: Itatiaia, 1983.
- DIAS, A. F. de S. Dissertação de Mestrado. *Comparação e Descrição de Parâmetros Acústicos do Canto de Volatinia Jacarina no Contexto de Seleção Sexual*. Brasília: Universidade de Brasília, 2009.
- FAGERLUND, S. Acoustics and physical models of bird sounds. *Seminar in acoustics, HUT, Laboratory of Acoustics and Audio Signal Processing*, 2004.
- FELFILI, J. M.; SOUSA-SILVA, J. C.; SCARIOT, A. *Cerrado: Biodiversidade, Ecologia e Conservação*. [S.l.]: MMA - Ministério do Meio Ambiente, 2005.
- HAYES, M. H. *Processamento Digital de Sinais*. São Paulo: Bookman, 2006.
- HAYKIN, S.; VEEN, B. V. *Sinais e Sistemas*. Porto Alegre: Bookman, 2001.
- JELINEK, F. *Statistical methods for speech recognition*. Cambridge: The Mit, 1998. 283 p. ISBN 0262100665.
- KIL, D. H.; SHIN, F. B. Pattern recognition and prediction with applications to signal characterization. *AIP Press, American Institute of Physics*, 1996.
- KOHN, A. F. *Reconhecimento de padrões: uma abordagem estatística*. Sao Paulo: Universidade de Sao Paulo, 1998. 205 p.
- KRAUSE, B. L. The niche hypothesis. *The Soundscape Newsletter 06*, June 1993.
- KWAN, C.; MEI, G.; ZHAO, X.; REN, Z.; XU, R.; STANFORD, V.; ROCHET, C.; AUBE, J.; HO, K. Bird classification algorithms: Theory and experimental results. *ICASSP2004*, v. 289-292, 2004.
- LEE, C.-H.; HAN, C.-C.; CHUANG, C.-C. Automatic classification of bird species from their sounds using two-dimensional cepstral coefficients. *IEEE Transactions on Audio*, v. 16, n. 8, p. 1541–1550, 2008.
- MARINI, M. A.; GARCIA, F. I. Conservação de aves no Brasil. *Megadiversidade*, v. 1, n. 1, 2005.
- MARQUES, A. B. Abordagens sobre a bioacústica na ornitologia - conceitos básicos. *Atualidade Ornitológicas OnLine*, Universidade Estadual do Norte Fluminense, Campos dos Goytacazes, 2008.
- MARQUES, A. B. Tese de Doutorado. *Avaliação do Canto do Trinca-Ferro (Saltator similis lafresnaye e d'orbigny 1837) em Relação ao Processo de Domesticação e suas Implicações na Conservação das Aves Canoras*. Campos dos Goytacazes: Universidade Estadual do Norte Fluminense, 2009.

- MATLAB. *MATLAB Documentation, Signal Processing Toolbox*. MATLAB, 2010. Disponível em: <<http://www.mathworks.com/help/toolbox/signal/f4-9659.html>>. Acesso em: 16/11/2010.
- MORRISON, M. L. Bird populations as indicators of environmental change. *Current ornithology*, Plenum Press, New York, 1986.
- OPPENHEIM, A. V.; SCHAFER, R. W. *Discrete-Time Signal Processing*. second. New Jersey: Prentice Hall, 1999. ISBN 0-13-754920-2.
- PAPAMICHALIS, P. E. *Practical Approaches to Speech Coding*. [S.l.]: Prentice Hall, 1987.
- PERLO, B. van. *A Field Guide to the Birds of Brazil*. [S.l.]: Oxford University Press, 2009. ISBN 978-0195301557.
- PICONE, J. W. Signal modeling techniques in speech recognition. *Proceedings of The IEEE*, v. 81, n. 9, p. 1215–1247, September 1993.
- PINHEIRO, R. T. Conservação de aves do cerrado: Conhecimento e perspectivas - o caso do tocantins. *Anais do VIII Congresso de Ecologia do Brasil*, Caxambu, MG, Setembro 2007.
- PLANET. *Planet of Birds*. Planet of Birds, 2012. Disponível em: <<http://www.planetofbirds.com/category/s-america/tinamiformes/tinamidae>>. Acesso em: 23/01/2012.
- RUSSEL, S.; NORVIG, P. *Inteligência Artificial*. São Paulo: Elsevier, 2004. ISBN 85-352-1177-2.
- SANCHES, D. *Interferência Eletromagnética*. [S.l.]: Interciencia, 2003. 124 p. ISBN 8571930848.
- SAVE. *Sociedade para Conservação das Aves do Brasil - Save Brasil*. Save Brasil, 2010. Disponível em: <<http://www.savebrasil.org.br/?q=content/por-que-aves>>. Acesso em: 27/02/2010.
- SCHURMANN, J. *Pattern classification: a unified view of statistical neural approaches*. New York: John Wiley Sons, 1996. ISBN 0471135348.
- SELIN, A.; TURUNEN, J.; TANTTU, J. T. Wavelets in recognition of bird sounds. *EURASIP Journal on Applied Signal Processing*, n. 2007, p. 1–9, 2007.
- SICK, H. *Ornitologia Brasileira*. Brasília: Universidade de Brasília, 1984.
- SICK, H. *Ornitologia Brasileira*. Rio de Janeiro: Ed. Nova Fronteira, 1997. 912 p.
- SILVA, J. M. C.; SANTOS, M. P. D. A importância relativa dos processos biogeográficos na formação da avifauna do cerrado e de outros biomas brasileiros. *Biodiversidade: Ecologia e Conservação do Cerrado*, 2005.
- SILVEIRA, L. F. Mundo das aves - os cantores que pouca gente vê. *Revista Cães e Cia*, v. 374, 2010. Disponível em: <<http://www.caes-e-cia.com.br/>>.
- SMITH, S. W. *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego: California Technical Publishing, 1999. ISBN 0-9660176-6-8.

SOMERVUO, P.; HÄRMÄ, A.; FAGERLUND, S. Acoustics and physical models of bird sounds. *IEEE Transactions on Audio, Speech and Language Processing*, v. 14, n. 6, November 2006.

SOUZA, D. *Aves do Brasil*. Belo Horizonte: Itatiaia, 1987.

STOICA, P.; MOSES, R. *Introduction to Spectral Analysis*. NJ: Prentice Hall, 2005.

WIKIAVES. *Tinamidae - Os Tinamídeos e o Homem*. Wikiaves, 2011. Disponível em: <<http://www.wikiaves.com.br/tinamidae>>. Acesso em: 02/11/2011.

WIKIAVES. *Wikiaves, a Enciclopédia das Aves do Brasil*. Wikiaves, 2011. Disponível em: <<http://www.wikiaves.com.br/>>. Acesso em: 08/04/2011.

WIKIPEDIA. *MP3*. wikipedia, 2011. Disponível em: <<http://pt.wikipedia.org/wiki/MP3>>. Acesso em: 03/11/2011.

XENO. *Xeno-canto - sharing bird songs from around the world*. Xeno-canto, 2011. Disponível em: <<http://www.xeno-canto.org/>>. Acesso em: 08/04/2011.

Lista de Artigos Publicados

CONCEIÇÃO, Paulo Francisco; MACHADO, Paulo César Miranda. **Abordagem para Extração Automática de Características para Reconhecimento de Aves Tinamidae Baseada em Análise Espectral**. Anais da 63ª Reunião Anual da SBPC – Cerrado: Água, Alimento e Energia. Goiânia – GO, 2011.

CONCEIÇÃO, Paulo Francisco; MACHADO, Paulo César Miranda; LEMOS, Rodrigo Pinto. **Sistema Automático para Segmentação e Determinação de Características de Sinais Sonoros de Aves do Cerrado Brasileiro a Partir da Análise Tempo-Frequência**. Anais do SBrT 2011 – XXIX Simpósio Brasileiro de Telecomunicações. Curitiba – PR, 2011.

CONCEIÇÃO, Paulo Francisco; MACHADO, Paulo César Miranda; CARVALHO, Fernando Aparecido. **Aplicação de Filtros Digitais em Sinais Sonoros de Aves para Sistema de Reconhecimento Automático**. Anais do SBrT 2011 – XXIX Simpósio Brasileiro de Telecomunicações. Curitiba – PR, 2011.

CONCEIÇÃO, Paulo Francisco; MACHADO, Paulo César Miranda; LEMOS, Rodrigo Pinto. **Automatic Recognition of Birds by Time-Frequency Analysis**. CLEI 2011 – XXXVII Conferência Latinoamericana de Informática. Quito, Equador, 2011.