

Pré-processamento de Sons para Reconhecimento Automático de Pássaros

Fernando Aparecido CARVALHO, Paulo César Miranda MACHADO

Escola de Engenharia Elétrica e de Computação, UFG, Goiânia-GO 74001970

Email: feapca@gmail.com, pcommachado@gmail.com

Palavras-chave: Reconhecimento de padrões, filtros digitais, segmentação, espectograma

INTRODUÇÃO

Com o avanço da tecnologia de reconhecimento de padrões, sistemas cada vez mais automáticos se tornam possíveis. “Um padrão é qualquer entidade da qual é possível extrair algum tipo de característica, seja ela simbólica ou numérica, e o que as técnicas computacionais de reconhecimento de padrões buscam, a partir destas características, é por uma maneira eficiente de organizarem estes padrões em agrupamentos ou classes que compartilhem determinadas semelhanças” (NOGUEIRA, 2006). Exemplos de reconhecimento de padrões são: identificação de impressão digital, reconhecimento óptico de caracteres, identificação de seqüências de DNA, reconhecimento de fala, reconhecimento de sons de aves e animais, etc.

O crescimento das cidades e a devastação do cerrado têm causado a migração de pássaros para regiões urbanizadas à procura de alimento. Muitos pássaros são encontrados em lixões a céu aberto ou até mesmo áreas alagadas próximas a aeroportos aumentando o risco de acidentes e incidentes com aeronaves. Com a degradação de ecossistemas, habitat de espécies nativas, o monitoramento dessas espécies por biólogos se torna necessário a fim de prever as possíveis conseqüências no meio-ambiente. O monitoramento pode ser feito por sistema automático de reconhecimento de som (vocalização) de aves.

Os sons dos pássaros são divididos em cantos e chamadas. Cantos são geralmente mais longos e complexos e estão relacionados à procriação e a defesa territorial, enquanto chamadas são mais curtas e estão relacionadas a alarme, vôo e alimentação (FAGERLUND, 2007).

Os métodos de reconhecimento de padrões são divididos em quatro fases: pré-processamento, extração de características, classificação, questões sobre classificação e pós-processamento. O pré-processamento, tratado neste trabalho, consiste na normalização do sinal, remoção de ruído e segmentação.

Revisado pelo orientador.

Orientando: Fernando Aparecido Carvalho. Orientador: Paulo César Miranda Machado

METODOLOGIA

Neste trabalho utilizaremos arquivos de sons coletados da internet (CORNELL LAB OF ORNITHOLOGY, WIKIAVES). Nosso enfoque será na Perdiz, ave da família dos Tinamídeos, por ter um som característico e que aparenta ser mais fácil de aplicar e adquirir domínio sobre as técnicas de pré-processamento. Como os arquivos de sons são coletados nos mais diversos formatos (*mp3*, *flv*, *ra*, etc), primeiramente convertem-se os arquivos para o mesmo formato, sendo escolhido o formato *wav*. Os arquivos coletados possuem, na sua maioria, a frequência de amostragem de 44100 Hz, o que é compatível com o espectro de som dos pássaros, que se situa entre 100 Hz e 8 kHz (KWAN, 2006).

O primeiro passo do pré-processamento é eliminar a componente contínua do sinal, calculando-se a média do sinal e subtraindo-a do próprio sinal a fim de se obter um sinal com média igual a zero. Em seguida, para que todos os sinais tenham a mesma intensidade, eles são normalizados isto é, cada sinal é dividido pelo seu valor máximo, obtendo-se um sinal normalizado entre -1 e 1 (KWAN, 2004), (SELIN, 2007).

Como os sons são normalmente gravados no ambiente em que as aves vivem, as gravações estão sujeitas aos mais variados tipos de ruídos, dentre os quais foram encontrados os seguintes tipos:

- Estalos (cliques): são pontos de saturação, ou seja, valores muito diferentes dos valores vizinhos em localizações aparentemente aleatórias, que podem ser percebidos nitidamente através de análise gráfica. Esse tipo de ruído pode ser observado na Fig. 1.

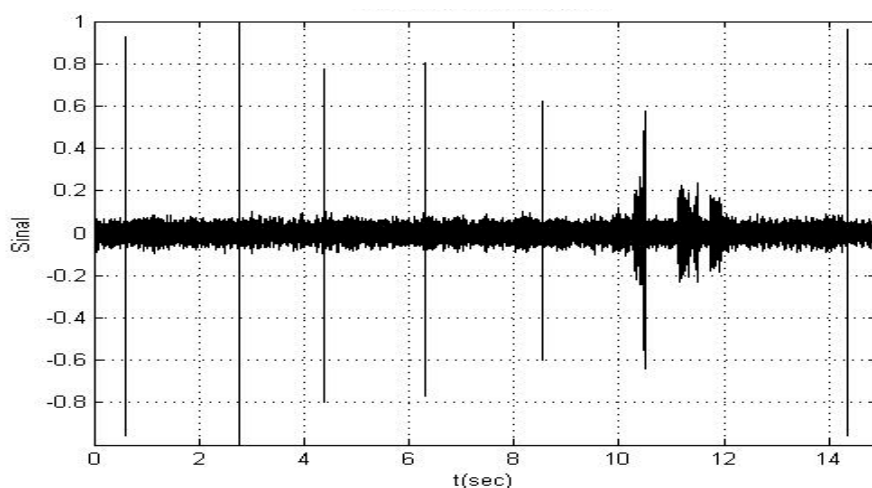


Figura1: Som de perdiz apresentando estalos.

- Chuva: é um som contínuo e o pingar é bastante intenso e rápido (Fig. 2). A interação é majoritariamente entre líquido e sólido o que gera maiores cavidades de ressonância. A energia sonora encontra-se distribuída por todas as frequências a partir de 500-700 Hz (DA SILVA, 2007).

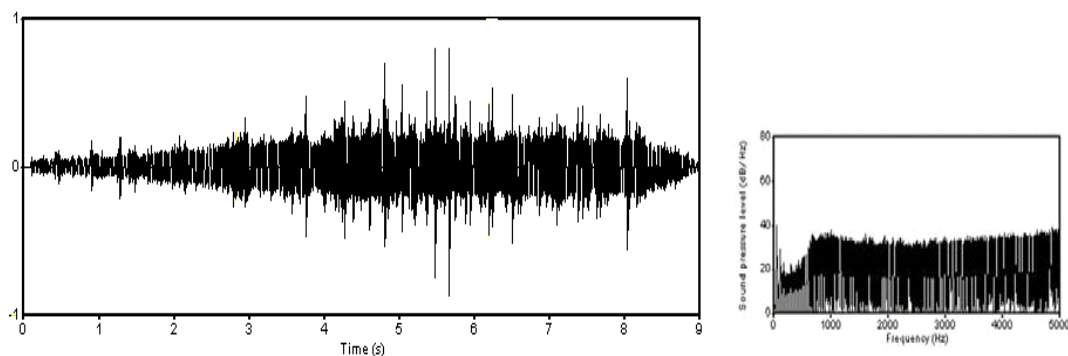


Figura 2: Forma de onda e espectro do som da chuva. Fonte: (DA SILVA, 2007).

- Grilos: os machos produzem um som roçando uma asa contra a outra (essas asas possuem uma série de pelos na borda) para atraírem a fêmea para a reprodução. Os machos produzem sons consecutivos de pequena duração cuja frequência é restrita a 2500 a 3000 Hz, conforme Fig. 3 (DA SILVA, 2007).

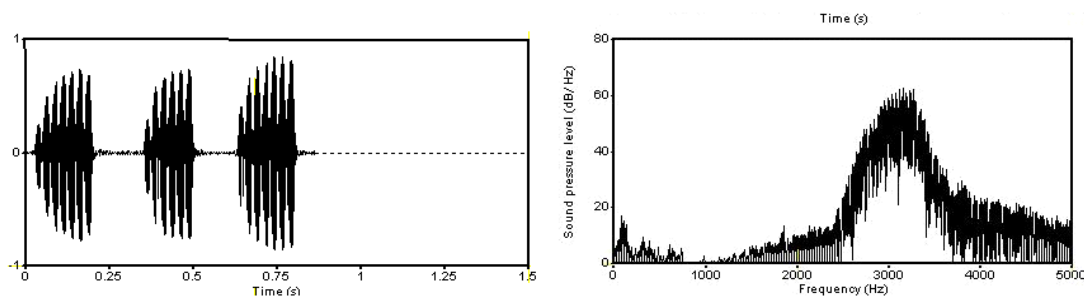


Figura 3: Forma de onda e espectro do som do grilo. Fonte: (DA SILVA, 2007).

- Vento: apresenta sinal contínuo com aumento gradual do sinal até valores mais elevados (Fig.4). Apresenta maior energia para frequências menores que 500 Hz, atingindo com menores taxas de energia cerca de 2000 Hz (DA SILVA, 2007).

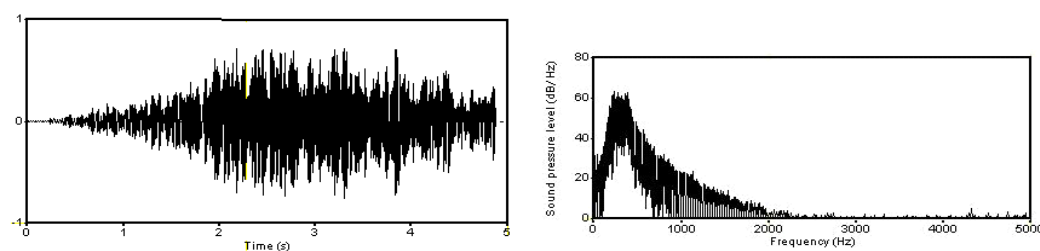


Figura 4: Forma de onda e espectro do som do vento. Fonte: (DA SILVA, 2007).

- 60 Hz e harmônicos: distorção do sinal apresentando altas taxas de energia para valores de frequência de 60 Hz e múltiplos (Fig. 5).

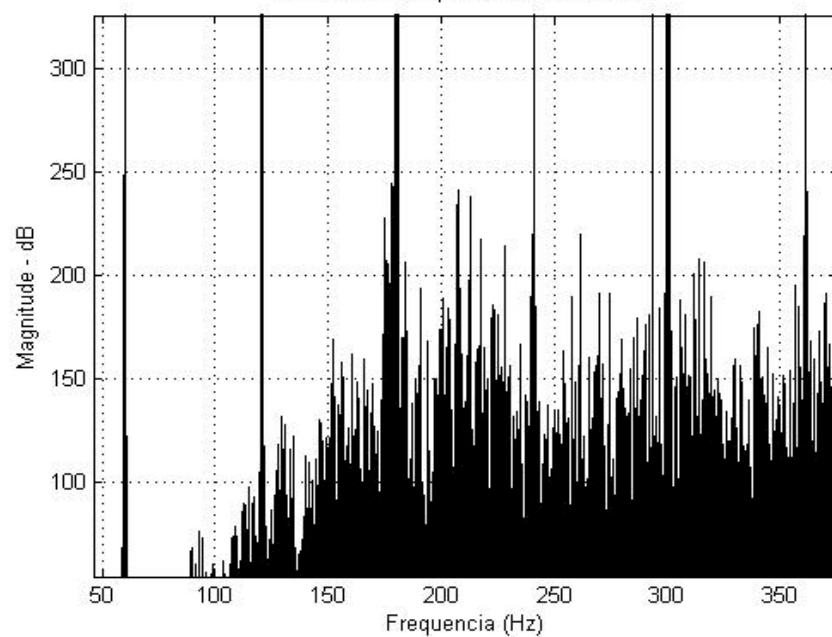


Figura 5: Espectro de som de perdiz apresentando harmônicos em 60 Hz e múltiplos.

O próximo passo do pré-processamento é a utilização de filtros para a eliminação dos ruídos. Os filtros utilizados foram Butterworth, Chebyshev do tipo 1 e 2, elíptico e Comb:

- Butterworth possui banda passante plana, porém, sua variação da magnitude é baixa sendo necessária altas ordens para implementação de certas especificações particulares.
- Chebyshev possui variação da magnitude maior que o Butterworth, mas possui ondulações (ripples) na banda passante ou na banda de bloqueio.
 - Chebyshev do tipo 1 apresenta ondulações na banda passante e grandes variações da magnitude a baixas ordens.
 - Chebyshev do tipo 2 apresenta ondulações na banda rejeitada e possui menor variação de magnitude que o do tipo 1.
- Elíptico permite maior taxa de variação da magnitude, porém, apresenta ondulações na banda passante e na banda de rejeição.
- Comb (pente) adiciona uma parte do sinal, ao próprio sinal, levemente defasado cancelando partes indesejadas do sinal.

A Fig. 6 mostra a resposta em frequência do filtro Comb:

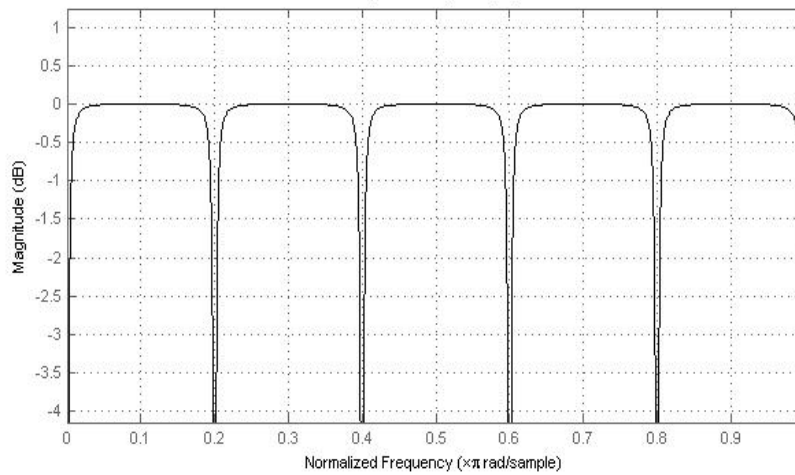


Figura6: Resposta em frequência do filtro Comb.

Fonte: <http://www.mathworks.com/help/toolbox/dsp/ref/fdesign.comb.html>

A Fig. 7 mostra a resposta em frequência dos demais filtros, todos com os mesmos parâmetros:

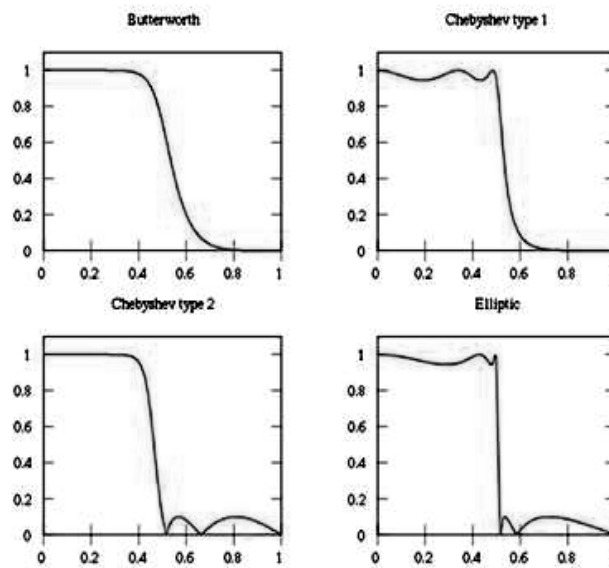


Figura 7: Resposta em frequência Butterworth, Chebyshev 1 e 2 e Elíptico.

Fonte: http://pt.wikipedia.org/wiki/Filtros_Butterworth

RESULTADOS E DISCUSSÃO

Primeiramente foi retirada a componente contínua, sinal foi normalizado e utilizado o filtro Comb para eliminar a interferência de 60 Hz e seus harmônicos da vocalização de uma perdiz (Fig. 8).

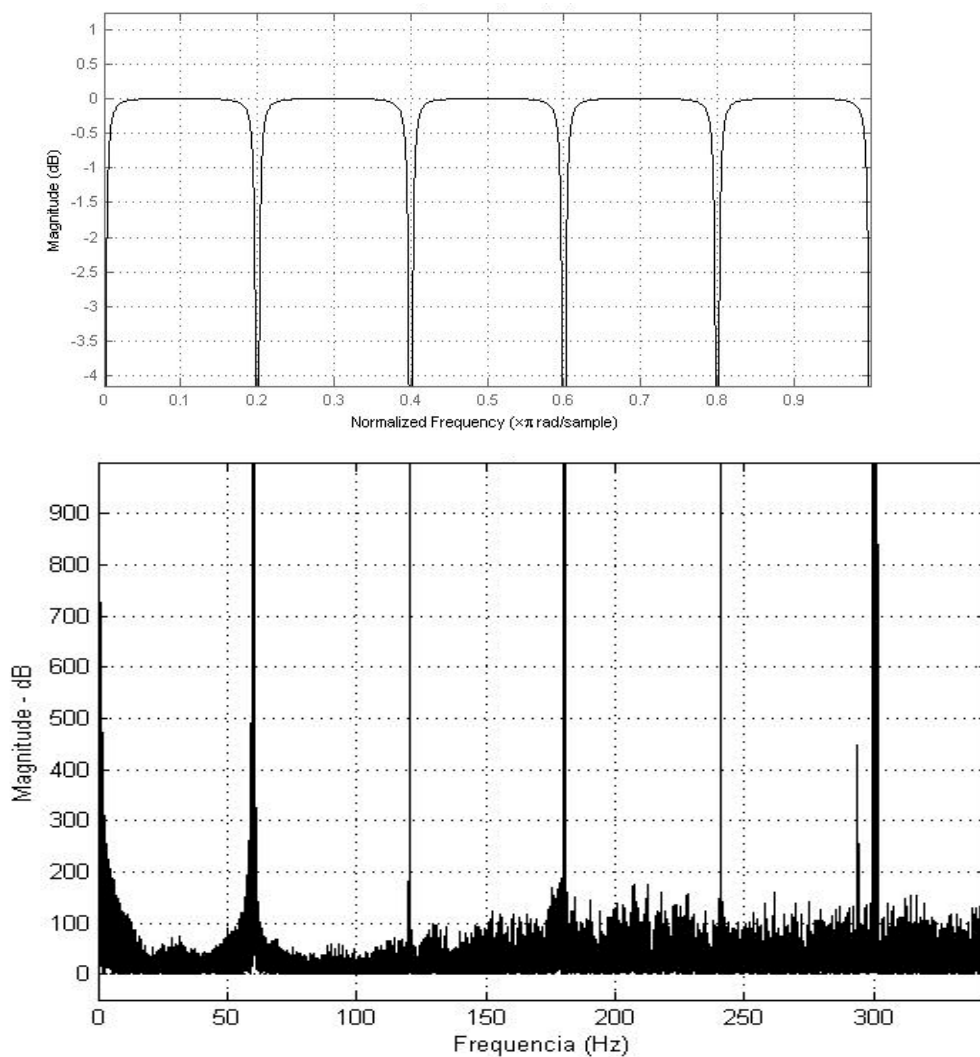


Figura 8: Filtro Comb aplicado a um sinal contendo interferência da frequência de 60 Hz e seus múltiplos (harmônicos).

A Fig. 9 mostra o sinal após a eliminação da interferência da frequência de 60 Hz e seus harmônicos, verificando-se que o sinal ainda apresenta bastante ruído. A mesma figura mostra também o espectrograma e o sinal no domínio da frequência.

A este sinal aplicou-se quatro tipos de filtros diferentes, todos com a mesma ordem (no caso ordem 6) para verificar o efeito dos vários tipos de filtros na eliminação de ruídos em sinais de cantos de pássaros.

As figuras de 10 a 13 mostram o sinal, seu espectrograma e o sinal no domínio da frequência após o uso dos filtros Butterworth, Chebyshev1, Chebyshev2 e Elíptico, respectivamente.

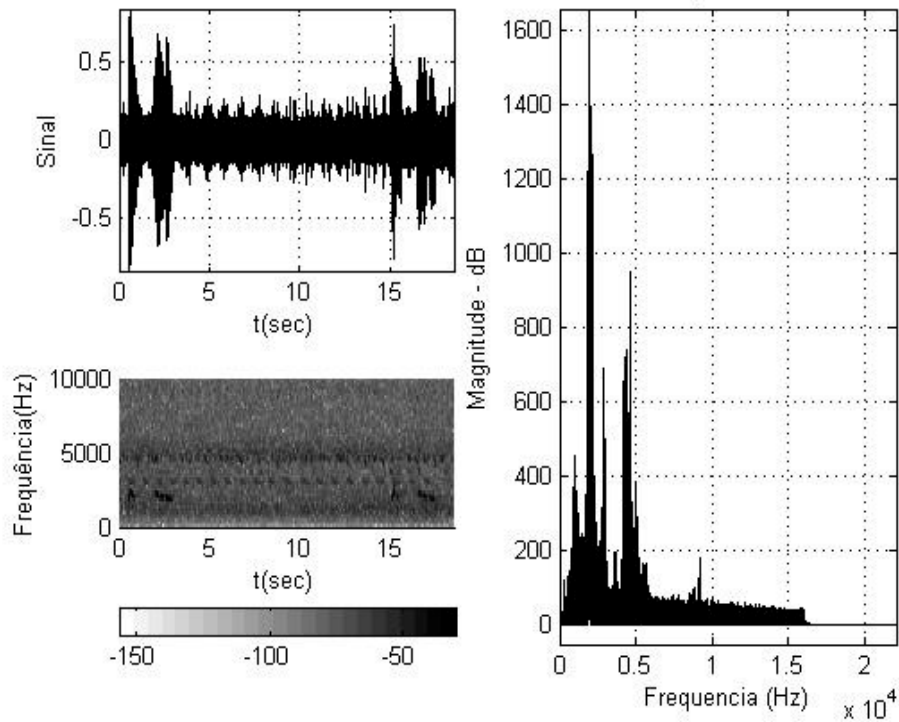


Figura 9: Sinal com ruído, após a eliminação da interferência da frequência de 60 Hz e seus harmônicos, seu espectrograma e sinal no domínio da frequência.

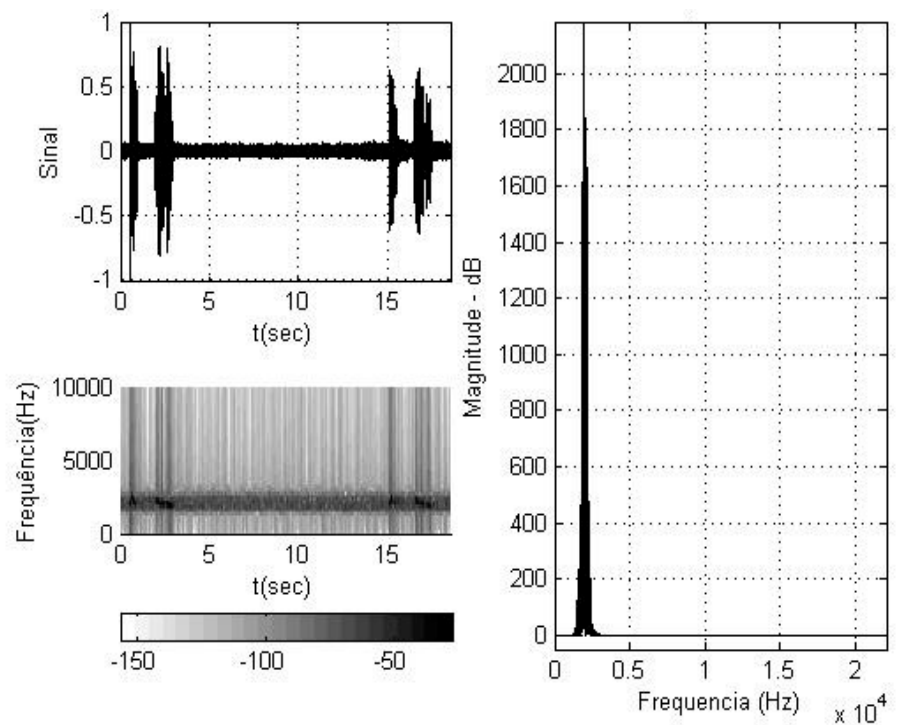


Figura 10: Sinal após utilização do filtro Butterworth (ordem 6), seu espectrograma e sinal no domínio da frequência.

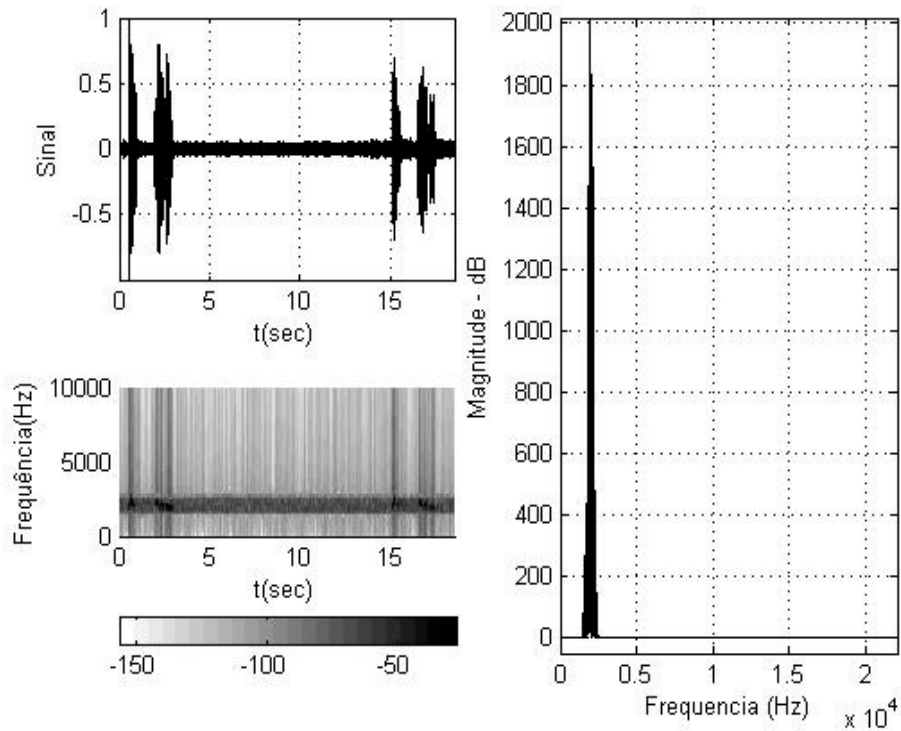


Figura 11: Sinal após utilização do filtro Chebyshev1 (ordem 6), seu espectrograma e sinal no domínio da frequência

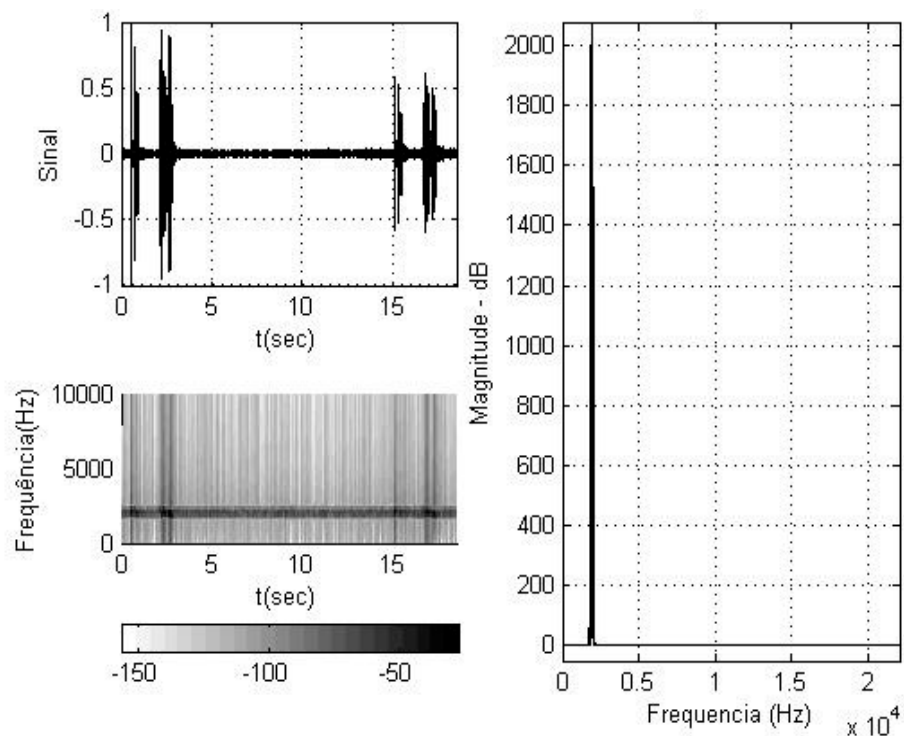


Figura 12: Sinal após utilização do filtro Chebyshev2 (ordem 6), seu espectrograma e sinal no domínio da frequência.

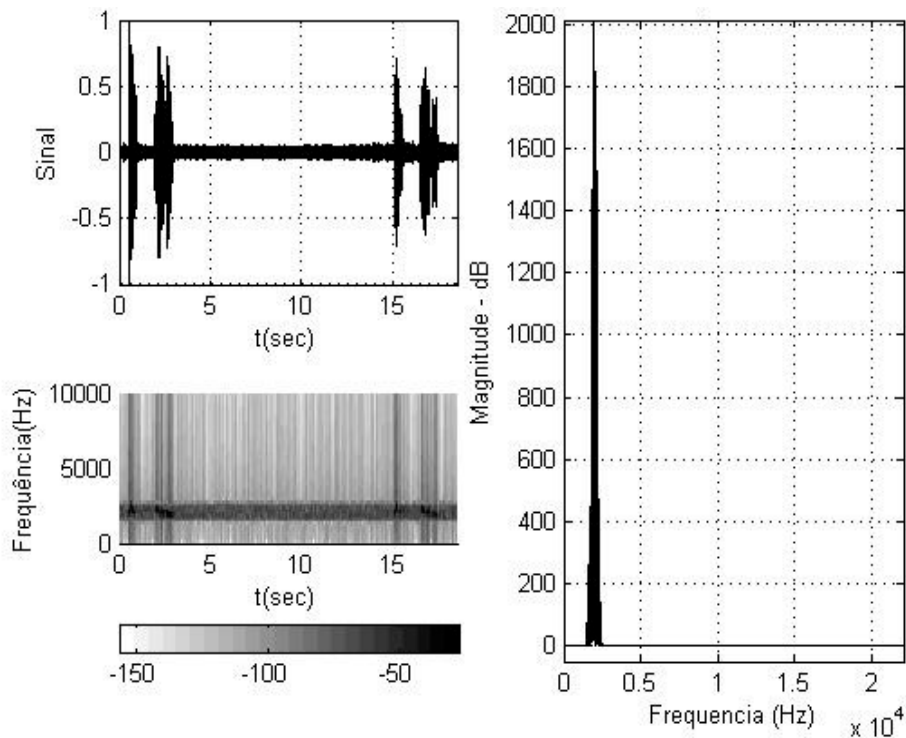


Figura 13: Sinal após utilização do filtro Elíptico (ordem 6), seu espectrograma e sinal no domínio da frequência.

Todos os filtros são do tipo passa banda onde a frequência de corte superior e inferior são iguais a 120% e 80% da frequência de maior energia, respectivamente.

Quando comparados os sinais das figuras de 10 a 13, resultantes do processo de filtragem, embora todos os filtros tenham conseguido deixar bem evidente o sinal do canto da ave, nota-se vantagem do filtro Chebyshev2, pois o sinal está bem mais nítido. Ao analisar os espectrogramas nota-se que o filtro Chebyshev2 realmente conseguiu clarear mais o sinal, deixando permanecer no espectro uma faixa de frequência bem próxima aos limites máximo e mínimo definidos como parâmetro do filtro passa banda.

Com base no espectrograma resultante da filtragem os valores da Densidade Espectral de Potência (PSD – Power Spectral Density) Tempo-Frequência são analisados com o 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

- $y[i]=1$, se $\max(P(:,i))$ for maior ou igual ao limiar;
- $y[i]=0$, se $\max(P(:,i))$ for menor que o limiar;

em que 0 representa silêncio, 1 representa som e $P(:,i)$ são as colunas da matriz de PSD.

Após o vetor y ser 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.

A figura 14 mostra o resultado da utilização da técnica para a segmentação do som de uma Perdiz com duas regiões de canto da ave. Percebe-se que neste caso a segmentação funcionou com estimativa bem precisa das regiões de som, mostrando claramente as duas regiões de canto da ave.

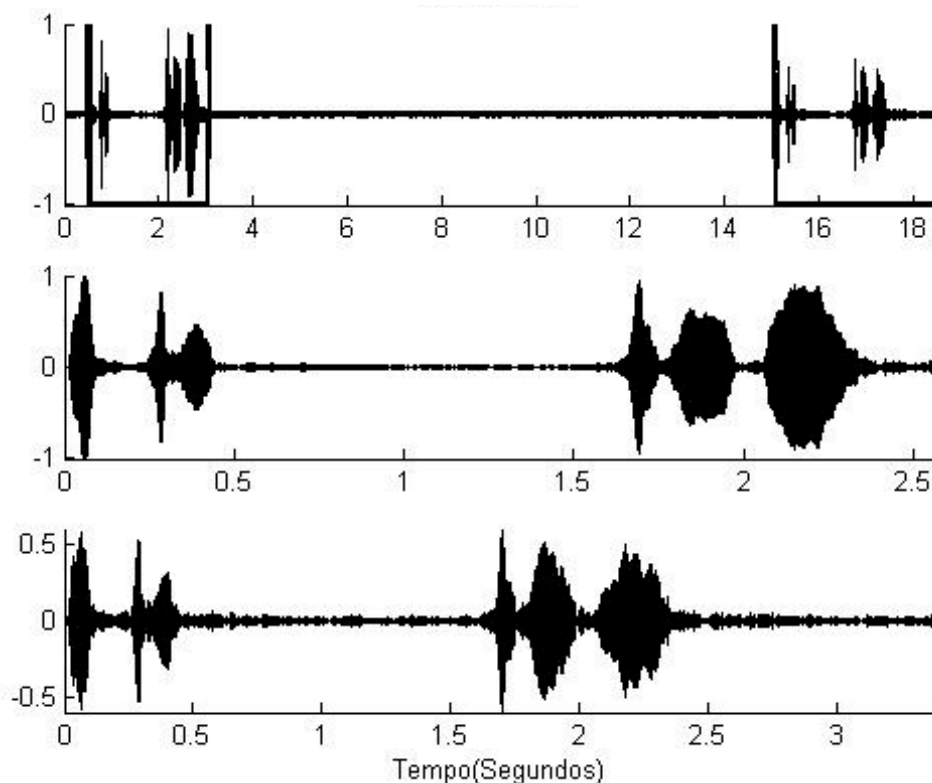


Figura 14: Sinal após filtragem e segmentação.

CONCLUSÕES

Neste trabalho foram apresentados procedimentos para pré-processamento de sinais de cantos de pássaros como a retirada da componente contínua, normalização do canto, eliminação da interferência de 60Hz e harmônicos e filtragem dos variados tipos de ruídos.

Assim, com os testes realizados, conclui-se que, de maneira geral, qualquer um dos filtros poderia ser escolhido como padrão para a etapa de pré-processamento, pois todos conseguiram eliminar os distúrbios e atenuar o ruído fora da faixa de frequência da ave. Entretanto, o filtro Chebyshev2 apresentou, levemente, um melhor desempenho, conseguindo clarear mais o espectro do sinal.

A metodologia utilizada para a segmentação do sinal mostrou-se eficiente, separando claramente as regiões de canto da ave das regiões de silêncio.

Como trabalho futuro pretende-se estudar e implementar métodos de extração de características para o reconhecimento automático de aves.

REFERÊNCIAS

DA SILVA, S. M. C. **Traços Acústicos e Perceptivos de Sons Não Verbais e da Fala, dissertação (mestrado em Ciências da Fala e da Audição)** - Escola Superior de Saúde da Universidade de Aveiro, Aveiro-Portugal, 2007.

NOGUEIRA, A.; AZEVEDO, J.; BAPTISTA, V.; SIQUEIRA, S. Um Overview Sobre Reconhecimento de Padrões, **VIII SEGeT**, pp. 1-11, 2006.

FAGERLUND, S.; Bird Species Recognition Using Support Vector Machines, **EURASIP Journal on Advances in Signal Processing**, pp. 1-11, 2007.

KWAN, C.; MEI, G.; ZHAO, X.; REN, Z.; XU, R.; STANFORD, V.; ROCHET, C.; AUBE, J.; HO, K. C. Bird Classification Algorithms: Theory and Experimental Results, **ICASSP2004**, pp. V289-V292, 2004.

C. Kwan et al., An automated Acoustic System to Monitor and Classify Birds, *Eurasip – Journal on Applied Signal Processing*, pp. 1-19, 2006.

CORNELL LAB OF ORNITHOLOGY. Disponível em <http://macaulaylibrary.org/>. Acesso em 08/04/2011.

WIKIAVES. Disponível em <http://www.wikiaves.com.br/>. Acesso em 08/04/2011.

SELIN, A.; TURUNEN, J.; TANTTU, J. T. Wavelets in Recognition of Bird Sounds, **EURASIP Journal on Applied Signal Processing**, vol. 2007, pp. 1-9, 2007.