

RECONHECIMENTO DE AVES DE NOMES ONOMATOPÉICOS

Célio Seixo de BRITO Junior, Paulo César Miranda MACHADO

Escola de Engenharia Elétrica e de Computação, UFG, 74001-970, Goiânia, GO
Celiojunior01@gmail.com, pcmmachado@gmail.com

Palavras-chave – Reconhecimento automático, filtro digital, detecção de frequência, entropia.

1 INTRODUÇÃO

Atualmente uma aplicação de Reconhecimento de Padrões [1,2] que vem sendo objeto de estudos é o monitoramento de aves através do reconhecimento automático de seus sons [3]. Na região centro-oeste do Brasil, com a crescente devastação do cerrado, alguns animais e aves estão em processo de extinção e hoje é possível encontrar na cidade animais e aves que só se encontravam nos campos, devido à perda de seu habitat natural. No sentido de preservação de espécies, um sistema de monitoramento e reconhecimento de pássaros pode ser utilizado para verificação das espécies existentes em uma região e para monitoramento de espécies ameaçadas de extinção [4].

O foco deste trabalho são as aves de nomes onomatopéicos, ou seja, aves cujos cantos característicos dão nome à espécie, como o Bem-te-vi, Fogo-apagou, Saracura-três-potes, Jaó e Tiziu. Algumas são aves que vivem em regiões de mata rasteira e densa, o que faz com que sejam mais escutadas do que vistas, tornando difícil o estudo de suas espécies (Jaó e Saracura-três-potes) ou são aves que tiveram que se adaptar ao grande avanço da urbanização (Bem-te-vi, Tiziu e Fogo - apagou) [5]. Essas aves foram escolhidas por serem aves típicas da fauna brasileira, que produzem sons bem característicos.

2 MÉTODO

Os sons utilizados neste trabalho foram coletados na internet (<http://www.wikiaves.com.br>, <http://www.xeno-canto.org/>) e, como foram gravados em condições diferentes, possuem diferentes níveis de intensidade. O primeiro passo no tratamento desses sons é a extração de seu valor médio, de maneira que todos sinais terão valor médio zero, e a sua normalização em relação ao seu valor máximo, com os valores dos sinais variando entre -1 e 1 (Ricke, 2005).

Neste trabalho as filtragens são realizadas de forma automática, ou seja, o próprio programa encontra a faixa de frequência na qual o filtro deve atuar e a ordem do mesmo. A detecção da faixa de frequência da vocalização das aves é feita automaticamente utilizando-se o espectrograma. Em seguida faz-se a filtragem do sinal. O filtro escolhido foi o Chebyshev tipo 2, por possuir a característica de ter a banda de passagem plana permitindo uma melhor captura do sinal desejado e ter uma banda de transição curta exigindo uma ordem menor para o filtro.

Após a filtragem é feita a segmentação do sinal para separar as regiões de som das regiões de silêncio, desenvolvida da seguinte maneira:

- A energia e o centróide são extraídos de todo o sinal de áudio;
- São estimados dois fatores de limiares para cada uma das características;
- Um critério de limite é aplicado sobre essas características para que possa ocorrer a separação entre o som e o silêncio.
- Por fim o som é detectado a partir dos critérios definidos anteriormente.

Utilizamos essas características pelo seguinte motivo: para sons onde o ruído de fundo não é muito elevado o valor da energia do som será maior que o do ruído e em situações onde o ruído é mais intenso os centróides espectrais do som serão maiores que os do ruído.

A energia do sinal é calculada pela equação (1):

$$E(i) = 1/N \sum_{n=1}^N |x_i(n)|^2 \quad (1)$$

O centróide é definido como o centro de gravidade do espectro do áudio. Este recurso é utilizado para trabalhar com características geométricas do espectro. Seu cálculo é feito pela equação (2).

$$C_i = \frac{\sum_{k=1}^N (k + 1) X_i(k)}{\sum_{k=1}^N X_i(k)} \quad (2)$$

Após o cálculo da energia e do centróide de cada janela, faz-se a separação da região de som da região de silêncio da seguinte maneira:

- Calcula-se o histograma dos valores da energia e centróide;
- Aplica-se um filtro simples de suavização do sinal;
- Detectam-se os máximos locais do histograma;

Neste trabalho o reconhecimento é feito a partir da frequência mais intensa e da entropia do sinal. A entropia é um recurso a ser utilizado se o sinal de áudio apresentar mudanças bruscas no nível de energia. A energia é normalizada com a equação (3), onde e_j é a seqüência normalizada da sub-janela de energia para cada janela e a entropia dessa seqüência é calculada a partir da equação (4):

$$e_j^2 = E_{sub-framej} / E_{framei} \quad (3)$$

$$Z(i) = - \sum_{j=1}^k e_j^2 \log_2(e_j^2) \quad (4)$$

3 RESULTADOS E DISCUSSÃO

Primeiramente os sinais foram normalizados entre -1 e 1, foram extraídos os valores médios dos sinais, de forma que a média dos sinais ficaram nulas, e foram filtrados, eliminando-se os ruídos. A Fig. 1 mostra um sinal, com muito ruído, e o resultado da filtragem com o filtro chebyshev2. Observa-se na figura que o filtro utilizado foi eficiente na eliminação dos ruídos e na preservação do canto da ave.

A Fig. 2(a) mostra o sinal após a realização da filtragem, o centróide e a energia do sinal. Pode-se observar que para esse sinal tanto a energia como o centróide são muito maiores na região do canto que na de silencio, o que torna possível realizar a segmentação.

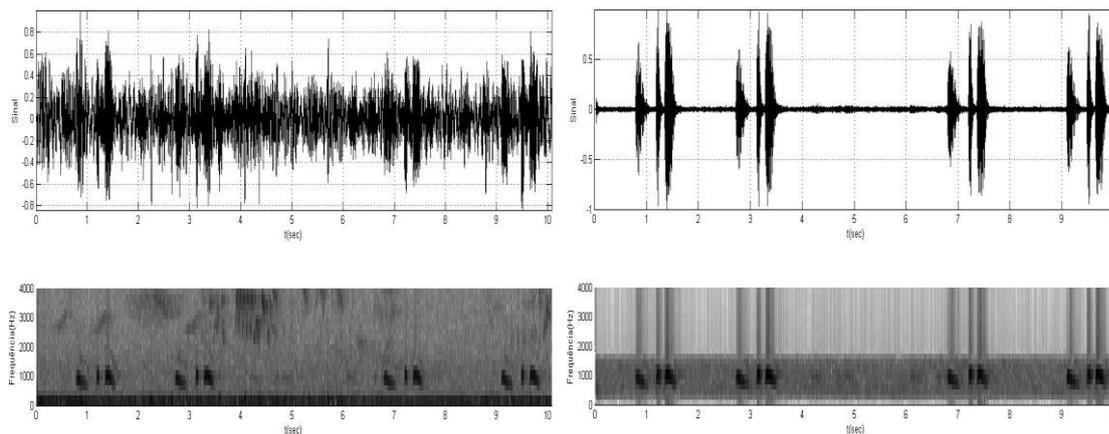


Figura 1. Sinal e espectrograma de um canto Fogo-apagou antes e depois da aplicação do filtro chebyshev tipo2.

A Fig. 2(b) mostra o sinal original e os sinais resultantes da segmentação, mostrando que o método utilizado foi eficiente em encontrar as quatro regiões em que há a vocalização da ave.

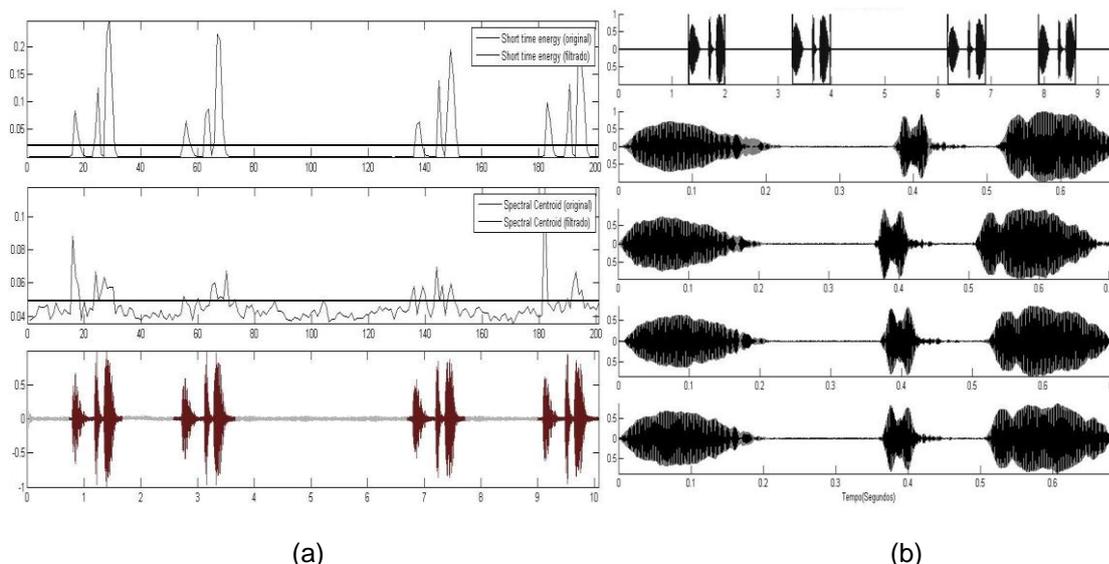


Figura 2. (a) Energia do sinal, centróide e sinal filtrado e (b) Regiões de som e silêncio detectadas automaticamente, mostrando quatro vocalizações da ave.

A fase de treinamento foi realizada com sons de 10 aves de cada espécie. Na fase de reconhecimento, foram utilizados os sons de 16 saracuras, 17 jaós, 16 tizius, 74 bem-te-vis e 42 fogo-apagou. Os resultados do reconhecimento são mostrados na tabela 1. Para as aves em estudo o reconhecimento foi satisfatório, pois 95% das aves foram reconhecidas corretamente.

Tabela 1. Resultados obtidos na fase de reconhecimento.

	Saracura	Jaó	Tiziu	Bem-te-vi	Fogo - Apagou
Saracura	16				
Jaó		17			
Tiziu			14	2	
Bem-te-vi			4	68	2
Fogo - Apagou					42

4 Conclusão

Neste trabalho foi proposta uma metodologia para o reconhecimento de aves de nomes onomatopéicos a partir de suas vocalizações. Foi utilizado um sistema de detecção automática da faixa de frequência para efeito de filtragem automática dos sinais, a segmentação do sinal em regiões de som e silêncio a partir

da energia e do centróide do sinal, e o reconhecimento utilizando como características a entropia do sinal e a frequência mais intensa do canto.

O método apresentado mostrou um bom desempenho para a maioria dos sinais testados, como mostram os resultados. A metodologia mostrou ter melhor resultado quando o sinal não possui um nível de ruído muito alto, mas mesmo nesses casos os resultados são satisfatórios. As características utilizadas para o reconhecimento (entropia e frequência de maior intensidade) se mostraram satisfatórias, já que mais de 90% das aves foram reconhecidas corretamente.

Como continuidade deste trabalho pretende-se desenvolver novas técnicas de eliminação de ruído, de segmentação e determinação de características do sinal, para efeito de reconhecimento automático de aves a partir da vocalização.

REFERÊNCIAS

- [1] R. O. Duda, P. E. Hart, D. G. Stork, ***Pattern classification (2nd edition)***, Wiley, New York, (2001).
- [2] S. Theodoridis. ***Pattern Recognition***. Academic Press, USA, (1999).
- [3] T. S. Brandes, ***Automated Sound Recording and Analysis Techniques for Birds Surveys and Conservation***, Bird Conservation International, p. 163-173, (2008).
- [4] S. Fagerlund, ***Automatic Recognition of Bird Species by Their Sounds***, Master Thesis, Dept. of Electrical and Communications Engineering, Helsinki University of Technology, (2004).
- [5] H. Sick, ***Ornitologia Brasileira - Uma Introdução***. Brasília: Ed. Universidade de Brasília, (1984).
- [6] A. D. Ricke, R. J. Povinelli and M. T. Johnson, ***Automatic Segmentation of Heart Sound Signals Using Hidden Markov Models***, ***Computers in Cardiology***, vol. 32, p. 953–956. (2005).
- [7] S. Theodoridis. ***Pattern Recognition***. Academic Press, USA, (1999).
- [8] D. H. Kil; F. B. Shin, ***Pattern Recognition and Prediction with Applications to Signal Characterization***, AIP Press, (1996).
- [9] J. W. Picone, ***Signal Modeling Techniques in Speech Recognition***, Proceedings of The IEEE, v81, n. 9, p. 1215-1247, 1993.