

Análise da possibilidade de deteção automática da vocalização de aves de rapina noturnas através da variação da pressão sonora

Vitor Rosão¹, Carlos Rosão², Inês Roque²,

¹SCHIU, Engenharia de Vibração e Ruído, Unip. Lda., Rua de Faro, Bloco B, 2.º F, 8005-463 Faro

vitorrosao@schiu.com

²shemahmforash@gmail.com

³LabOr - Laboratório de Ornitologia, Departamento de Biologia & Instituto de Ciências Agrárias e Ambientais Mediterrânicas, Universidade de Évora 7002 – 544 Évora, Portugal

iroque@uevora.pt

Resumo

Pretende-se com o presente artigo analisar a viabilidade de deteção automática da vocalização de aves de rapina noturnas através da variação da pressão sonora, com sonómetro ou equivalente, na medida em que este equipamento é de utilização generalizada e efetua de forma eficaz, rápida e fiável, o tratamento da informação acústica, permitindo assim a simplificação do algoritmo de deteção automática e, subsequentemente, uma maior rapidez de análise. Acresce ao referido o custo relativamente reduzido do processamento de sinal associado à determinação da variação da pressão sonora, custo este que, sendo demonstrada a viabilidade de deteção automática com o uso deste tipo de equipamento, como é objetivo do presente artigo, permitirá assim também uma redução significativa do custo usual da instrumentação para a deteção automática de indivíduos, em particular aves de rapina noturnas, grupo alvo do presente Estudo.

Palavras-chave: Deteção Sonora; Vocalização Animal; Aves de Rapina Noturnas; Variação da Pressão Sonora; Sonómetros.

Abstract

This article intends to analyze the feasibility of automatic detection of vocalizations of owls by means of the variation of sound pressure, using sound level meter or equivalent. The sound level meters are widely used and perform efficiently, rapidly and reliably the treatment of acoustical information, thereby enabling the simplification of the algorithm for automatic detection and subsequently increasing the speed of analysis. In addition to that, the relative low cost of the signal processing associated with the determination of the variation of sound pressure implies the possibility of significant reduction for the usual costs of instrumentation for the automatic detection of individuals, particularly owls, the group targeted with this study.

Keywords: Sound Detection; Animal Vocalization; Owls; Variations of Sound Pressure, Sound Level Meters.

PACS no. 43.80.Ka, 43.60.Bf, 43.58.Fm

1 Introdução

O estudo das aves de rapina noturnas ou estrigiformes (comummente conhecidas como mochos e corujas) envolve vários constrangimentos logísticos, dadas as características comportamentais do grupo, constituído sobretudo por espécies com hábitos crepusculares e noturnos. O método de censo mais expedito para caracterização destas comunidades implica a realização de pontos de escuta, sendo os contactos quase exclusivamente auditivos. Desta forma, o conhecimento das vocalizações características de cada espécie é fundamental para a sua identificação.

O método dos pontos de escuta implica a seleção de um conjunto de estações de amostragem onde um observador permanece durante um período de tempo determinado, registando todas as aves detetadas [1].

Uma vez que as vocalizações das estrigiformes são determinadas geneticamente, têm uma grande importância taxonómica. Adicionalmente existem muito poucas variações entre dialetos, indivíduos ou subespécies. Dentro deste grupo, cada vocalização tem um determinado significado, como p. ex. a defesa do território ou atração do parceiro [2]. É durante o período reprodutor que estas aves estão mais ativas vocalmente, sendo portanto nessa fase do ciclo de vida que a realização de pontos de escuta é mais eficaz.

Existem em Portugal sete espécies de estrigiformes das quais cinco são residentes. As espécies mais comuns são o mocho-galego *Athene noctua* e a coruja-do-mato *Strix aluco*. O mocho-galego é uma ave de pequeno porte (~23 cm) que pode ser observada durante o dia, ocupando uma grande diversidade de habitats de Norte a Sul do país. É particularmente abundante em olivais, montados de azinho e paisagens em mosaico (parcelas agrícolas intercaladas com sebes), evitando áreas florestais fechadas onde a coruja-do-mato é abundante. A coruja-do-mato é uma ave de médio porte (~40 cm) tipicamente florestal, ocupando preferencialmente povoamentos extensos de carvalhos, sobreiros, azinheiras, castanheiros e pinheiros [3].

As vocalizações têm um papel importante na estruturação das comunidades de estrigiformes, funcionando como sinais de comunicação intra e interespecífica. O mocho-galego é frequentemente atacado por outras aves de rapina: num estudo de Zuberogoitia *et al.* [4] a espécie foi a que sofreu maior número de ataques de entre todas as estrigiformes europeias, sendo as principais espécies atacantes a coruja-das-torres *Tyto alba* e a coruja-do-mato. Assim, é de esperar que evite áreas ocupadas por estas e outras estrigiformes com maior capacidade de competição, embora haja casos de coexistência pacífica, sobretudo se o alimento estiver disponível em abundância e se houver baixa disponibilidade de locais de nidificação [5].

No caso do mocho-galego as vocalizações da espécie poderão estar na origem de atração conspecífica ou induzir o estabelecimento de pares reprodutores como forma de defesa contra a competição conspecífica; por outro lado, funcionarão como sinal de segurança para o mocho-de-orelhas *Otus scops*, a menor estrigiforme nidificante em Portugal (~20 cm), a qual mostrou igualmente preferência por nidificar em áreas ocupadas pelo mocho-galego [6].

1.1 Contributo acústico

Tendo em conta o descrito, assume especial relevância a capacidade de deteção automática de vocalizações de estrigiformes.

Uma vez que as características da vocalização variam de espécie para espécie, e dentro da mesma espécie podem variar em função da idade, do sexo e da finalidade da vocalização, não se enquadra nos objetivos deste estudo inicial, procurar desenvolver um algoritmo capaz de distinguir automaticamente todas essas características. Procura-se sim, dentro da perspectiva de uma primeira abordagem à possibilidade de utilização das variações da Pressão Sonora para a deteção automática das vocalizações, a possibilidade de gravação áudio prolongada e registo simultâneo e sincronizado da variação da pressão sonora, sem presença humana, e a posterior deteção automática – através do algoritmo desenvolvido – dos momentos de maior probabilidade de ocorrência de vocalizações, a serem confirmados através de audição humana, o que, apesar de corresponder a uma primeira abordagem, terá a grande vantagem de poder diminuir significativamente o tempo de audição humana do ficheiro áudio prolongado para deteção das vocalizações de interesse. De qualquer forma, aborda-se, tendo em conta os resultados obtidos, a possibilidade do desenvolvimento futuro de algoritmos baseados na variação da Pressão Sonora para deteção mais criteriosa das vocalizações de interesse.

São conhecidos estudos mais detalhados de deteção automática das vocalizações [7-11], estando inclusive disponível no mercado instrumentação específica, contudo tais sistemas são significativamente dispendiosos, o que talvez justifique a não utilização em Portugal, e em outros países [12], de tais sistemas, o que justifica também este estudo, no sentido da procura de sistemas mais económicos de deteção automática das vocalizações.

O presente estudo centraliza-se em duas gravações áudio, uma de vocalizações da coruja-do-mato, e outra de vocalizações de mocho-galego, e nos respetivos registos das variações dos Níveis de Pressão Sonora através de sonómetro, e no desenvolvimento de algoritmo de deteção automática dos momentos de ocorrência das vocalizações.

2 Gravações áudio e respetiva variação dos Níveis de Pressão Sonora

No site <http://www.schiu.com/corujas> encontram-se disponíveis as duas gravações alvo de análise, denominadas respetivamente por:

1. “CorujaDoMato.mp3” “MochoGalego.mp3”

Nos gráficos seguintes apresentam-se as respetivas variações dos Níveis de Pressão Sonora.

A análise dos gráficos permite verificar o seguinte:

1. A variação dos Níveis de Pressão Sonora aparenta de fato poder ser um método eficaz de deteção automática dos momentos de ocorrência das vocalizações, possibilitando assim uma diminuição significativa do tempo de análise do ficheiro áudio, uma vez que é expectável que os momentos de ocorrência de picos de Nível de Pressão Sonora sejam os momentos de ocorrência de vocalizações.
2. Quanto maior for o ruído de fundo e/ou quanto menor for a magnitude das vocalizações, em termos de Níveis de Pressão Sonora, mais difícil será a deteção das vocalizações através deste método. Apesar de corresponder a uma limitação do método, essa limitação ocorre também no

método de auscultação humana *in situ*, onde é necessário selecionar locais onde o ruído de fundo seja reduzido e onde seja mais provável uma audição perceptível das vocalizações.

3. A ocorrência de picos de Nível de Pressão Sonora associados a eventos que não correspondam às vocalizações de interesse vai fazer com que o método detete momentos de provável vocalização que não correspondem na realidade a vocalizações. Contudo, uma adequada seleção do local de “audição” deverá minimizar significativamente estas ocorrências sendo inclusive possível aperfeiçoar o algoritmo no sentido de descartar automaticamente as “falsas” vocalizações, como se explicará no capítulo 4.

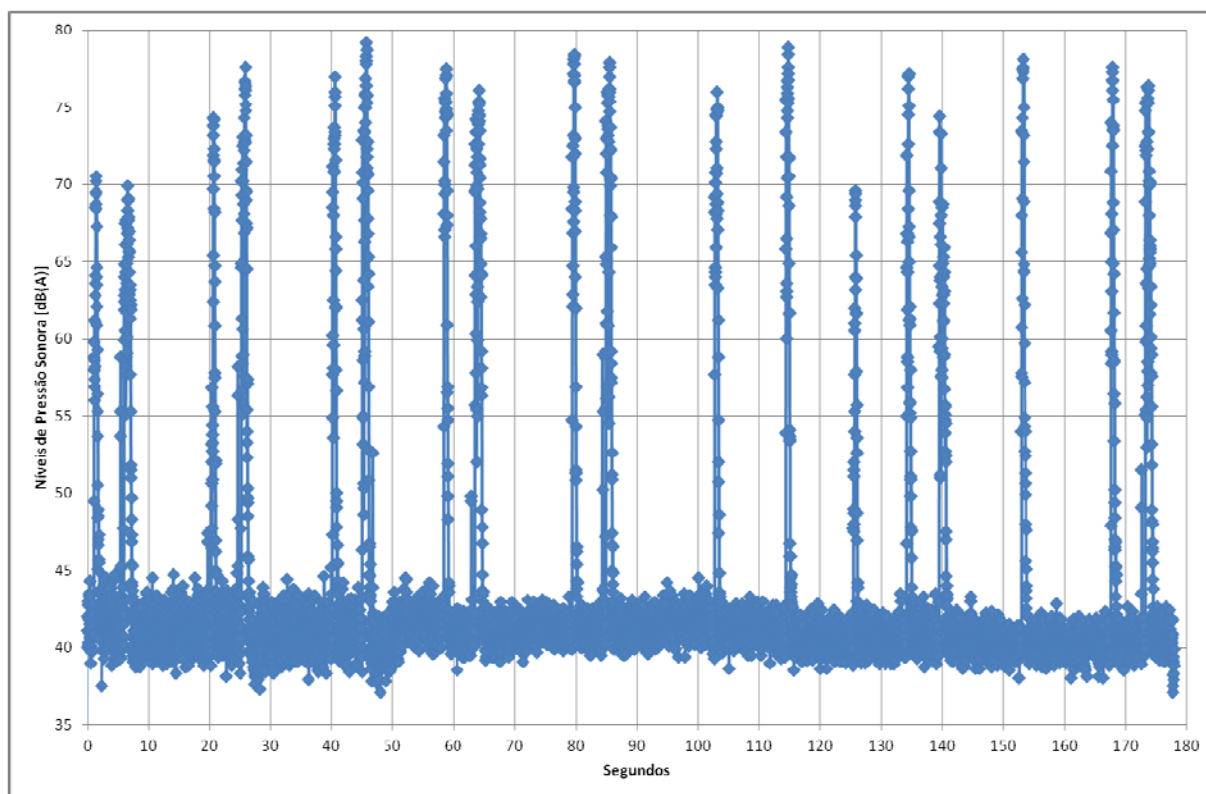


Figura 1 – Variação dos Níveis de pressão Sonora na gravação da coruja-do-mato

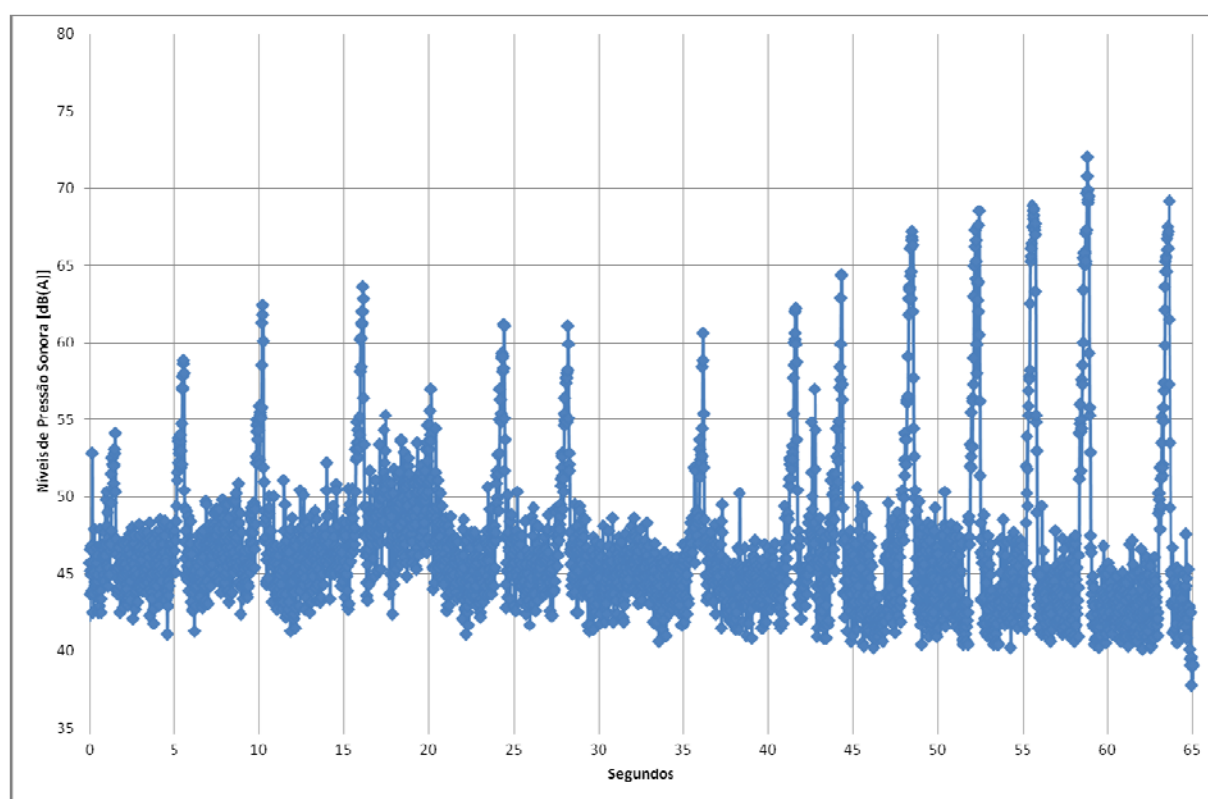


Figura 2 – Variação dos Níveis de pressão Sonora na gravação do mocho-galego

3 Algoritmo desenvolvido

Tal como é possível ver nas Figuras 1 e 2, a variação dos Níveis de Pressão Sonora apresenta máximos bem localizados com algum tipo de variabilidade devido a ruído de fundo. De forma a detectar os máximos, o algoritmo desenvolvido utiliza diversos passos: Pós-Processamento, Limiar e Selecção de Picos, baseados na literatura de detecção de Onsets [13].

3.1 Pós Processamento

O passo de pós processamento está relacionado com a simplificação dos passos seguintes de Limiar e Selecção de Picos. Desta forma, consiste na aplicação de métodos de normalização e filtros aos dados que se está a analisar, dando origem à função que iremos chamar $d(n)$.

3.2 Limiar

Mesmo após o pós-processamento, ainda existem máximos que não correspondem àquilo que queremos obter mas sim a ruído apenas. Desta forma, é comum a criação de uma Função de Limiar que separa os máximos que interessam daqueles que não [13,14]

Como primeira aproximação pode-se criar uma Função de Limiar constante [15], porém, em sinais de áudio dinâmicos este tipo de função não obtém resultados satisfatórios [13]. Para ultrapassar este problema, é comum a utilização de Funções de Limiar dinâmicas [14].

Existem diversos tipos de Funções Limiar dinâmicas, porém aquele que tendencialmente obtém melhores resultados é construído através da mediana local dos dados pós-processados [15] e representa-se de seguida por $f(n)$:

$$f(n) = \delta + \lambda \text{median}(|d(n - M)|, \dots, |d(n + M)|) \quad (1)$$

, onde δ , λ e M são parâmetros ajustáveis.

3.3 Seleção de Picos

Após os processamentos anteriores, a escolha dos máximos que correspondem às corujas está reduzida à identificação de máximos locais dos dados pós-processados que estejam acima da Função de Limiar:

$$c(n) = \begin{cases} 1 & \text{if } d(n) > f(n) \\ & \text{and } d(n - 1) \leq d(n) \leq d(n + 1) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

onde $c(n)$ é a função que representa onde existem sons de corujas.

3.4 Algoritmo

O algoritmo, desenvolvido de acordo com o discutido nos três subcapítulos anteriores, foi incluído num *framework* geral de áudio desenvolvido em [16] e que está disponível de forma livre e gratuita em [17].

Os parâmetros discutidos nos subcapítulos anteriores podem ser ajustáveis de forma a melhorar os resultados. Para uma deteção automática o mais fiável possível, será necessária uma amostra significativa de vocalizações de corujas de forma a ser utilizada para obtenção dos valores padrão para o conjunto de parâmetros do algoritmo.

4 Possibilidades de evolução

O sistema ideal de deteção automática das vocalizações deveria ser capaz de distinguir os diferentes tipos de vocalizações, se se trata do mesmo ou de outro indivíduo e se se trata efetivamente de vocalizações ou de outro tipo de evento, por exemplo a passagem de veículos rodoviários.

De referir que no método atual de audição humana *in situ* é necessário que a pessoa envolvida na audição tenha alguma experiência, pois é preciso distinguir os diferentes tipos de vocalizações e associá-los a uma determinada espécie e/ou indivíduo, através da perceção da eventual movimentação da ave e/ou da sua localização, tentando perceber, através dessa eventual movimentação ou localização, se se trata de apenas um indivíduo ou de mais do que um indivíduo. Conjugando a experiência do observador com as variações que ocorrem nos Níveis de Pressão Sonora nos diferentes posicionamentos e movimentações dos indivíduos, afigura-se, à partida, ser possível desenvolver um

algoritmo, com base nas variações dos Níveis de Pressão Sonora, que aponte direções relativamente ao tipo de vocalização, à sua localização (eventualmente com o uso de mais do que um ponto de escuta no mesmo local, sincronizados, para verificação das diferenças de tempos e de pressão sonora no sentido de determinar a localização da origem da vocalização) e se se trata apenas de um indivíduo ou de mais do que um indivíduo.

De forma a evidenciar a possibilidade de evolução do método, ilustram-se nos gráficos seguintes as variações dos Níveis de Pressão Sonora associados aos diferentes tipos de vocalização detetados nas gravações áudio:

1. Primeira parte de frase da coruja-do-mato, ocorrente aos segundos 1, 20, 40, 59, 80 (1:20), 103 (1:43), 115 (1:55), 126 (2:06), 135 (2:15), 153 (2:33), 168 (2:48) do ficheiro “CorujaDoMato.mp3”, apresentando-se dois exemplos da variação dos respectivos Níveis de Pressão Sonora na Figura 3.
2. Segunda parte de frase da coruja-do-mato, ocorrente aos segundos 6, 25, 45, 64 (1:04), 85 (1:25), 140 (2:20), 174 (2:54) do ficheiro “CorujaDoMato.mp3”, apresentando-se dois exemplos da variação dos respectivos Níveis de Pressão Sonora na Figura 4.
3. Vocalização do mocho-galego, ocorrente aos segundos 2, 5, 10, 16, 20, 24, 28, 36, 41, 44, 48, 52, 55, 58, 63 (1:03) do ficheiro “MochoGalego.mp3”, apresentando-se dois exemplos da variação dos respectivos Níveis de Pressão Sonora na Figura 5.

Apresenta-se também (Figura 6) a variação típica dos Níveis de Pressão Sonora à passagem de veículos rodoviários a diferentes velocidades de acordo com a referência [18].

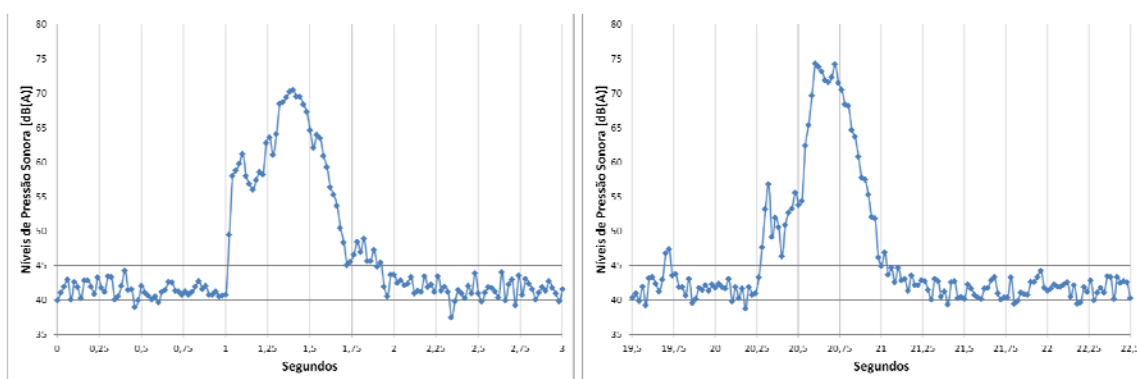


Figura 3 – Variação dos Níveis de Pressão Sonora aos segundos 1 e 20 da gravação “CorujaDoMato.mp3”

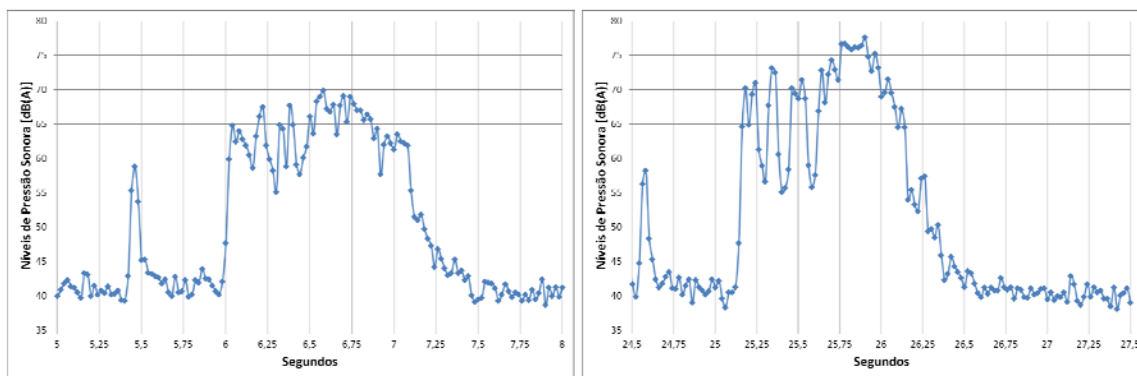


Figura 4 – Variação dos Níveis de Pressão Sonora aos segundos 6 e 25 da gravação “CorujaDoMato.mp3”

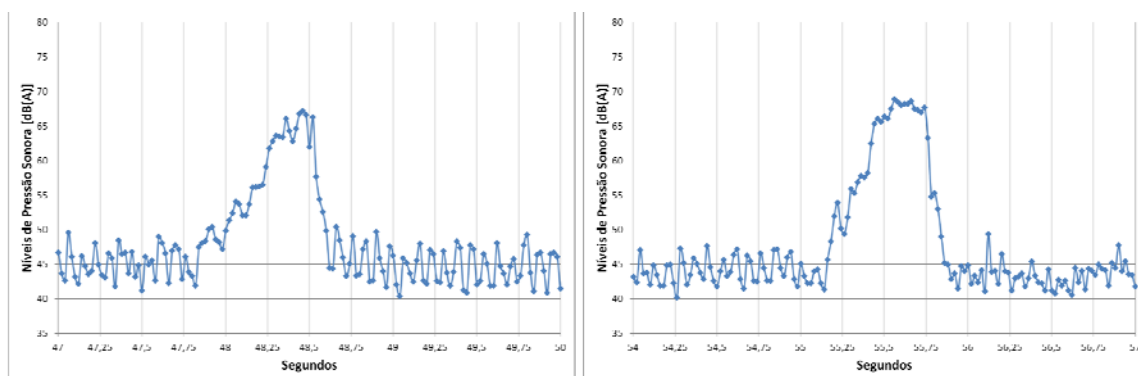


Figura 5 – Variação dos Níveis de Pressão Sonora aos segundos 48 e 55 da gravação “MochoGalego.mp3”

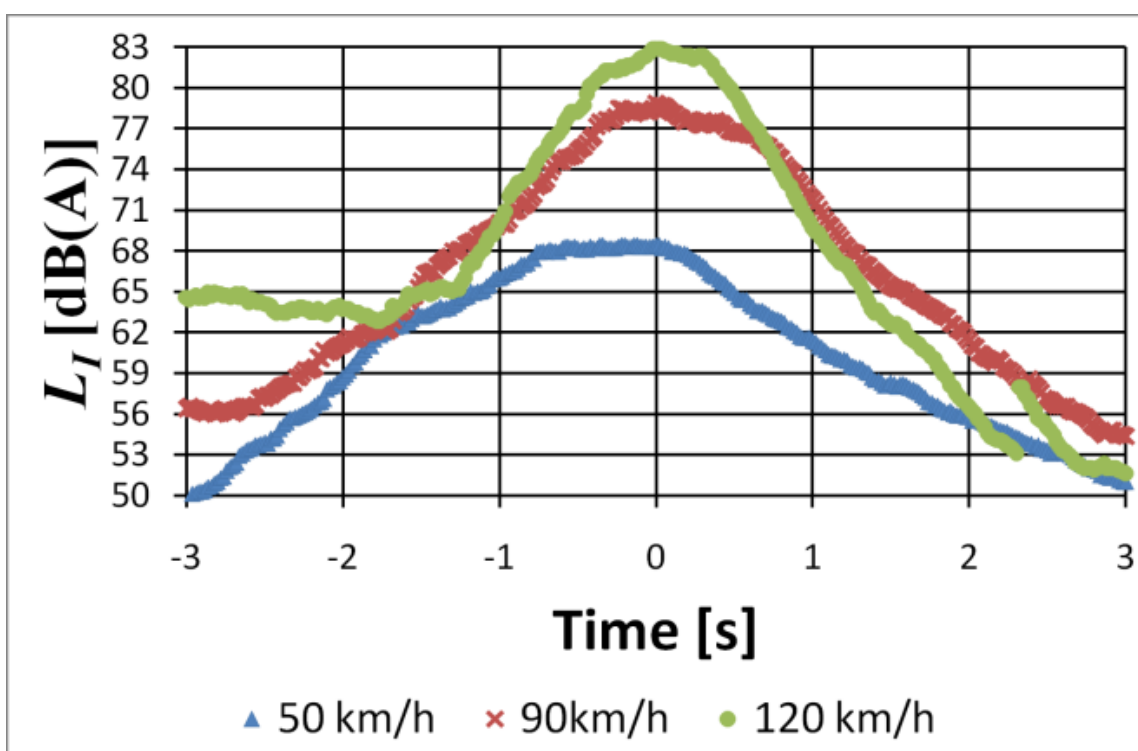


Figura 6 – Variação típica dos Níveis de pressão Sonora à passagem de veículos rodoviários [18]

Analisando os gráficos anteriores verifica-se o seguinte:

1. É possível distinguir um padrão distinto entre a variação dos Níveis de Pressão Sonora para a Figura 3 e para a Figura 4, o que significa a possibilidade de desenvolvimento de um algoritmo que consiga distinguir, através da variação dos Níveis de Pressão Sonora, os dois tipos de vocalização da coruja-do-mato detetados na gravação “CorujaDoMato.mp3”.
2. Aparenta ser mais difícil a distinção entre a vocalização da coruja-do-mato retratada na Figura 3 e a vocalização do mocho-galego retratada na Figura 5, o que faz antever a

necessidade de análise mais aprofundada destes tipos de variação, de forma a encontrar padrões distintos, ainda que seja perceptível uma subida mais continuada na Figura 5 do que na Figura 3, antevendo assim também a possibilidade de desenvolvimento de um algoritmo que consiga distinguir, através da variação dos Níveis de Pressão Sonora, as vocalizações da Figura 3 e da Figura 5.

3. O descartar de eventos não desejados aparenta ser simples no caso da passagem de veículos rodoviários, pois tendo em conta as variações apresentadas na Figura 6, as variações dos Níveis de Pressão Sonora são mais continuadas e demoram significativamente mais tempo a ocorrer (pelo menos 2 segundos, no caso da passagem de um veículo a 120 km/h) do que as vocalizações registadas nas duas gravações em análise (no máximo 1 segundo de ocorrência).
4. O descartar de outro tipo de eventos terá de ser analisado caso a caso e corresponderá sempre a uma hipótese de evolução do método tendo por base a experiência da sua implementação.

5 Conclusões

Considera-se que o algoritmo “simples” desenvolvido, ainda que corresponda a um primeiro desenvolvimento associado à análise da possibilidade de utilização das variações dos Níveis de Pressão Sonora na deteção automática de vocalizações, permite uma significativa poupança de tempo na análise dos ficheiros áudio e possui grande potencial de evolução, conforme se procurou evidenciar.

Referências

- [1] Bibby, C.J.; et al. – *Bird Census Techniques*. Second Edition. Academic Press, 2000.
- [2] König, Claus; Weick, Friedman – *Owls of the world*. Second Edition. A&C Black Publishers, 2008.
- [3] Cattr, Paulo; Costa, Helder; Elias, Gonçalo; Matias, Rafael – *Aves de Portugal: ornitologia do território continental*. Assírio & Alvim, 2010.
- [4] Zuberogoitia, Iñigo; et al. – *Interspecific aggression and nest-site competition in a European owl community*. Journal of Raptor Research, Vol. 39, pp 156-159, 2005
- [5] van Nieuwenhuyse, Dries; Génot, Jean-Claude; Johnson, David – *The Little Owl: Conservation, ecology and Behavior of Athene noctua*. Cambridge University Press, 2008.
- [6] Parejo, Deseada; Avilés, Jesus M., Rodríguez, Juan - *Alarm calls modulate the spatial structure of a breeding owl community*. Proceedings of the Royal Society Biological Sciences, published online before print January 25, 2012. doi: 10.1098/rspb.2011.2601
- [7] Brandes, T. Scott – *Automated sound recording and analysis techniques for bird surveys and conservation*. Bird Conservation International, Vol. 18, Supplement S1, pp S163-S173, September 2008.
- [8] Venier, Lisa A.; et. al. – *Evaluation of an automated recording device for monitoring forest birds*. Wildlife Society Bulletin, Vol. 36, Issue 1, pp 30–39, March 2012.
- [9] Jancovic, Peter; Kokuer, Munevver – *Automatic Detection and Recognition of Tonal Bird Sounds in Noisy Environments*. EURASIP Journal on Advances in Signal Processing, Article ID 982936, Vol. 2011.
- [10] <http://www.google.com/patents/US20080223307>

- [11] Briggs, Forrest – *Acoustic classification of multiple simultaneous bird species: A multi-instance multi-label approach*. J. Acoust. Soc. Am. 131 (6), June 2012.
- [12] MacLeod, Catriona J. – *Monitoring widespread and common bird species on New Zealand's conservation lands: a pilot study*. New Zealand Journal of Ecology 36(3), 2012.
- [13] J. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler - *A tutorial on onset detection in music signals*. IEEE Transactions on Speech and Audio Processing, 13(5):1035–1047, 2005.
- [14] S. Dixon - *Onset Detection Revisited*. In Proc. of the Int. Conf. on Digital Audio Effects (DAFx-06), pages 133–137, Sept. 2006.
- [15] I. Kauppinen. *Methods for detecting impulsive noise in speech and audio signals*. In 2002 14th International Conference on Digital Signal Processing Proceedings. DSP 2002 (Cat. No.02TH8628), volume 2, pages 967–970. IEEE.
- [16] Rosão, Carlos. (2012). *Onset Detection in Music Signals*. (Master's thesis)
- [17] Audio Framework – <https://github.com/Shemahmforash/Audio-Framework>
- [18] Rosão, Vitor; Conceição, Eusébio; Hazyova, Lucia – *Method to Determine the Speed of Vehicles by Means of Noise Levels Variation*. Lisboa, InterNoise 2010.