

Identificação Automática de Espécies de Pássaros a partir da Análise do Canto

Marcelo T. Lopes¹, Celso A. A. Kaestner¹, Carlos N. Silla Jr.², Alessandro L. Koerich³

¹Departamento de Informática, Universidade Tecnológica Federal do Paraná
Curitiba, PR, Brasil

²School of Computing, University of Kent
Canterbury, Kent, UK

³Programa de Pós-Graduação em Informática, Pontifícia Universidade Católica do Paraná
Curitiba, PR, Brasil

teiderlopes@gmail.com celsokaestner@utfpr.edu.br cns2@kent.ac.uk alekoe@computer.org

Abstract. *This paper deals with the automated bird species identification problem in which it is necessary to identify the species of a bird from its audio recorded song. Different features sets which summarize in different aspects the audio properties of the audio signal are evaluated in this paper together with machine learning algorithms. Experiments conducted in a dataset of three bird species have shown that it is possible to obtain very promising results.*

Resumo. *Este artigo trata do problema da identificação automática de espécies de pássaros onde busca-se identificar as espécies a partir de amostras gravadas do canto. São avaliadas diferentes características que podem ser utilizadas para modelar diferentes aspectos do sinal de áudio juntamente com algoritmos de aprendizagem de máquina. Experimentos realizados em um conjunto de dados de amostras de cantos de três espécies de pássaros mostraram que é possível obter resultados promissores a partir dos conjuntos de características utilizados.*

1. Introdução

Questões ecológicas ocupam atualmente o centro das discussões e preocupações da sociedade contemporânea. Para uma melhor compreensão e avaliação de nosso meio-ambiente, é necessário obter continuamente informações confiáveis a respeito da população de animais selvagens. Neste contexto, os pássaros tem um papel importante, pois eles são uma das mais numerosas classes que tem um contato direto com os seres humanos. Pássaros podem ser monitorados de maneira indireta a partir da gravação de seus sons, o que facilita significativamente o processo de monitoramento. Portanto, a utilização de métodos automáticos para a identificação de espécies de pássaros, incluindo tanto hardware quanto software, é fundamental para um monitoramento efetivo da quantidade e diversidade de pássaros presentes em um ecossistema específico [1, 2].

Diversos pesquisadores vêm tratando o problema da identificação automática de espécies de pássaros através de técnicas de reconhecimento de padrões. Dentre os inúmeros desafios para desenvolver um sistema capaz de monitorar pássaros de maneira indireta a partir de seu canto, uma tarefa fundamental é a identificação automática das espécies de pássaros. Neste problema, o canto de pássaros de espécies conhecidas, gravado em seus ecossistemas, é a informação disponível para servir como entrada de

um sistema de identificação automática. Cada um dos cantos gravados pode ser analisado para produzir um conjunto de características discriminantes que representam o canto do pássaro, de modo que seja possível caracterizar sua espécie. Portanto, se uma base de dados de canto de pássaros pré-definidos estiver disponível, é possível projetar métodos de classificação automática. A Figura 1 apresenta uma visão geral do método proposto para tal fim.

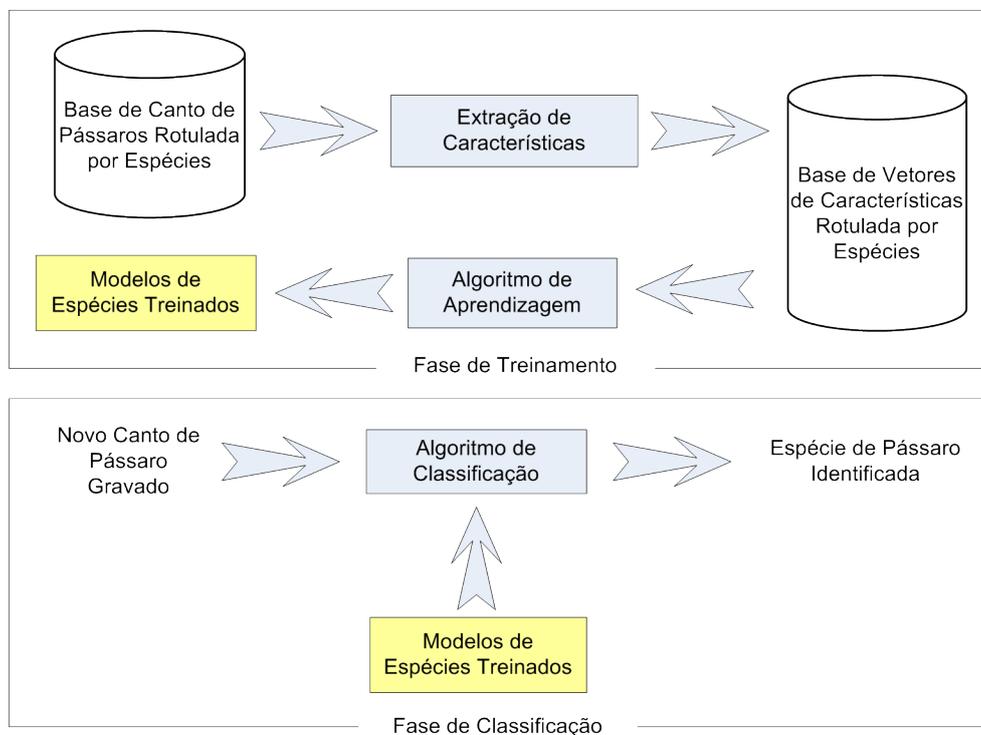


Figura 1: Visão geral do esquema de classificação de espécies de pássaros

Como na maioria dos problemas de reconhecimento de padrões, a definição do conjunto de características e a escolha do algoritmo de classificação são cruciais para um bom desempenho. Assim, este artigo apresenta um estudo inicial que compara diferentes características e classificadores que foram previamente empregados para a classificação de gêneros musicais, no problema da identificação de espécies de pássaros [3].

2. Identificação da Espécies de Pássaros

O problema da identificação automática da espécie de pássaros pode ser definido como a tarefa de encontrar a espécie de um pássaro específico a partir da análise do áudio gravado de seu canto. Os sons emitidos por pássaros podem ser classificados em cantos e chamados. Restringimos o escopo de nosso trabalho aos sons dos pássaros produzidos na forma de canto. Para representarmos os cantos, três conjuntos de características são empregados: MARSYAS, IOIHC e Sound Ruler.

MARSYAS: originalmente proposto por Tzanetakis e Cook [4] para o problema da classificação de gêneros musicais, pode ser dividido em três subconjuntos: características do conteúdo rítmico, características descrevendo a textura timbral e características relacionadas ao tom. O vetor de características final possui 64 dimensões.

IOIHC: os coeficientes do histograma de intervalos Inset-Onset estão relacionados às propriedades rítmicas do sinal musical, mais especificamente ao tempo entre o início dos pontos de ataque de eventos sucessivos ou notas [5]. O conjunto de características é computado a partir de uma função particular de periodicidade rítmica que representa

Tabela 1: F-Score sobre o conjunto de dados com canto de pássaros(%)

Algoritmo de Classificação	Conjunto de Características		
	Sound Ruler	IOIHC	MARSYAS
Naïve Bayes	99,7	43,5	86,9
3-NN	96,8	57,4	98,4
J48	99,0	61,0	99,7
MLP	98,7	68,0	99,7
SVM(Polinomial)	97,8	53,5	99,4
SVM (Pearson)	99,4	64,3	99,4

a saliência normalizada com respeito ao período dos intervalos inset-onset presentes no sinal. O conjunto de características final é representado por um vetor de características de 40 dimensões.

Sound Ruler: é uma ferramenta para medida e representação gráfica que pode calcular 36 características de um sinal de áudio pulsado, como tempo do pico, pico do pulso relativo, forma e duração dos pulsos, intervalos e duração dos pulsos, razão cíclica, razão entre o pico e a média do sinal, diversas medidas de energia em subintervalos, frequências fundamental, mínima e máxima e amplitudes relativas. Uma descrição detalhada destas características pode ser encontrada em [6].

Para a tarefa de classificação dos empregamos o algoritmo Naïve Bayes, o algoritmo dos três vizinhos mais próximos, árvore de decisão, uma rede neural do tipo *perceptron* multicamadas treinada com o algoritmo retropropagação e um classificador baseado em máquinas de vetor de suporte usando o algoritmo de otimização sequencial mínimo com as funções de kernel polinomial e Pearson IV.

3. Experimentos

Os experimentos foram realizados a partir de cantos de pássaros obtidos no website Xeno-Canto [7]. Os experimentos foram limitados somente a três espécies de pássaros: *Taraba major* (32 amostras), *Cercomacra tyrannina* (34 amostras) e *Thamnophilus doliatus* (35 amostras). Os registros foram obtidos diretamente de ambientes reais e incluem ruídos ambientais e sons de outros animais. Os registros foram segmentados em pulsos e deles extraímos os três conjuntos de características, gerando vetores com 64, 40 e 36 dimensões. A Tabela 1 apresenta a média ponderada sobre as três espécies para os diferentes algoritmos de classificação considerando validação cruzada 10-fold. Notamos que o conjunto IOIHC é o que fornece os piores resultados provavelmente devido ao fato de estarmos lidando com cantos de pássaros onde o conteúdo rítmico não diferente muito dentre as espécies. Por outro lado, os conjuntos extraídos pelo Sound Ruler e MARSYAS fornecem resultados interessantes.

Foi empregado o teste de Friedman com o procedimento estatístico post-hoc Shaffer para a comparação de múltiplos classificadores. O valor p do teste estatístico e o valor crítico corrigido foram de 0,0024 e 0,0166, 0,0303 e 0,0500 e 0,3864 e 0,0500 para IOIHC vs. MARSYAS, Sound Ruler vs. IOIHC e Sound Ruler vs. MARSYAS respectivamente. Considerando um nível de confiança de 95%, não há diferença significativa entre os resultados com os conjuntos MARSYAS e do Sound Ruler.

4. Conclusões

Os resultados mostraram que tanto os conjuntos de características obtidos a partir do MARSYAS e do Sound Ruler oferecem um bom desempenho na tarefa de identificação

automática da espécie de pássaros a partir da análise do canto dos pássaros para quase todos os algoritmos de classificação. Por outro lado, o conjunto de características IOIHC parece inadequado, pois ele explora as características rítmicas do canto dos pássaros, que são claramente diferentes do encontrado em músicas. Os testes estatísticos realizados mostram que as diferenças nos valores resultantes do F-score não são significativas entre as características geradas pelo MARSYAS e pelo Sound Ruler. Isso deve-se ao fato da maioria das características em ambos os conjuntos possuírem uma natureza similar, consistindo principalmente de valores dos MFCCs calculados em intervalos de tempo específicos.

Na literatura corrente verificamos que os diversos artigos tratam do problema da identificação automática de espécies de pássaros através da análise utilizam vários conjuntos de características e diferentes algoritmos de aprendizagem de máquina. Entretanto, nenhum deles se preocupou em fazer uma comparação direta sobre um conjunto de dados de referência, assim como não há um indício de qual conjunto de características é o mais adequado para representar os cantos de pássaros.

De modo a comparar os resultados obtidos neste artigo com os da literatura, destacamos que o nosso melhor resultado (99,7%) é superior a maioria dos resultados apresentados em artigos recentes. Entretanto, a comparação mais justa é com o trabalho de Vilches et al. [8] (98,4%), que utilizou as mesmas espécies de pássaros e o conjunto de características do Sound Ruler.

Referências

- [1] R. Bardeli, D. Wolff, F. Kurth, M. Koch, K-H. Tauchert and K-H. Frommolt, “Detecting Bird Songs in a Complex Acoustic Environment and Application to Bioacoustic Monitoring”, *Patt Recog Letters*, Vol.31, pp.1524-1534, 2010.
- [2] T.S. Brandes, “Automated Sound Recording and Analysis Techniques for Bird Surveys and Conservation”, *Bird Cons Intl*, Vol.18, pp.163-173, 2008.
- [3] T. Lidy, C. N. Silla Jr., O. Cornelis, F. Gouyon, A. Rauber, C. A. A. Kaestner and A. L. Koerich. On the Suitability of State-of-the-Art Music Information Retrieval Methods for Analyzing, Categorizing, Structuring and Accessing Non-Western and Ethnic Music Collections. *Signal Processing*, Vol.90, pp.1032-1048, 2010.
- [4] G. Tzanetakis and P. Cook, “Musical Genre Classification of Audio Signals”, *IEEE Trans. Speech Audio Process.*, Vol.10, pp.293-302, 2002.
- [5] F. Gouyon, S. Dixon, E. Pampalk and G. Widmer, “Evaluating Rhythmic Descriptions for Music Genre Classification”, *Proc. 25th Int’l AES Conf. on Virtual, Synthetic and Entertainment Audio*, London, UK, 2004.
- [6] Sound Ruler Web Site, disponível em <<http://soundruler.sourceforge.net/>> acessado em 24 de Junho de 2010.
- [7] Xeno-Canto Web Site, disponível em <<http://xeno-canto.org/>> acessado em 26 de Junho de 2010.
- [8] E. Vilches, I.A. Escolbar, E.E. Vallejo and C.E. Taylor, “Data Mining Applied to Acoustic Bird Species Recognition”, *Proc. 18th IEEE Intl Conf on Patt Recognition*, Hong Kong, China, pp.400-403, 2006.