

# BIOMETRIA DE INSETOS: Análise do sinal de áudio

Pablo Alves Pinheiro<sup>1</sup>, Marcelo Gonzaga de Oliveira Parada<sup>2</sup>  
<sup>1,2</sup> Departamento de Engenharia Elétrica, Centro Universitário FEI  
 pablo.tecno@hotmail.com; mparada@fei.edu.br

**Resumo:** O projeto consiste na identificação de insetos através da análise do sinal de áudio produzido por um inseto durante o voo. Foram utilizadas ferramentas para extração de parâmetros no domínio da frequência, como MFCC (*Mel-Frequency Cepstral Coefficients*) assim como bibliotecas para modelagem dos parâmetros baseados em misturas Gaussianas, como GMM (*Gaussian Mixture Model*). Os códigos utilizados foram desenvolvidos nos programas MATLAB® bem como Python com utilização do Hardware Raspberry Pi. A comparação dos resultados é apresentada através de gráficos e tabelas.

## 1. Introdução

O projeto tem como foco principal a identificação biométrica através do áudio capturado de diferentes insetos, utilizando técnicas atualmente empregadas para identificação de locutor, ou seja, identificação de seres humanos através do sinal de fala [1]. Assim como os seres humanos possuem diferenças em seu trato vocal que produzem sons com características acústicas diferentes, o mesmo ocorre em insetos que possuem diferentes tamanhos e tipos de asas e, portanto, produzem diferentes sons em frequências diferentes. Estas diferenças no áudio emitidos pelos insetos pode ser identificada a partir de uma análise em frequência, e a partir desses modelos matemáticos específicos de cada inseto podem ser extraídos para posteriormente utilizar como padrão para a identificação de insetos de mesma espécie.

O processo de identificação pode ser dividido em duas partes principais, elas são: *front-end* e *back-end* [2]. O *front-end* consiste em três etapas: filtragem do som, extração dos coeficientes mel-cepstrais (MFCC) e extração dos dados no formato HTK, conforme Figura 1.

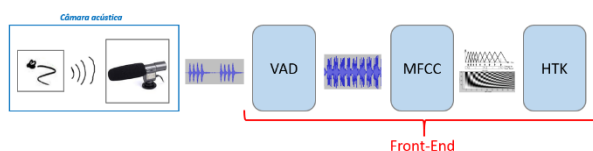


Figura 1 - Front-end.

- I. VAD (*Voice Activity Detector*)
- II. MFCC (*Mel-frequency cepstral coefficients*)
- III. HTK (*Hidden Markov Model Toolkit*)

O *back-end* é responsável por cadastrar um modelo e comparar este modelo com um indivíduo a ser analisado, através de técnicas de modelagem, uma das ferramentas mais usadas é o *Gaussian Mixture Model* (GMM) [3]. A comparação com o modelo já armazenado é realizada por métricas matemáticas como o *Log-Likelihood Ratio*

(LLR). A sequência de aplicação do *Back-end* na fase de utilização do sistema pode ser visualizada na Figura 2.

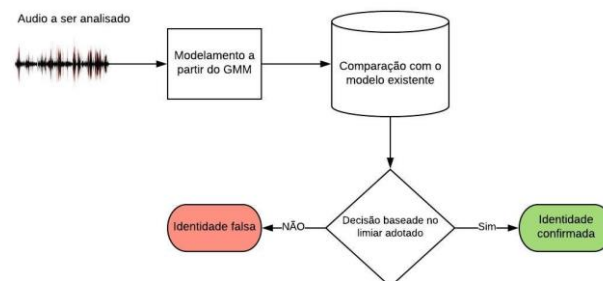


Figura 2 - Back-end.

## 2. Metodologia

Os testes realizados nesta iniciação científica foram divididos em duas etapas: testes baseados no MATLAB®, e os testes realizados em Python utilizando o Raspberry.

Na primeira etapa do projeto foi utilizado o software MATLAB® em conjunto com os Toolbox Voicebox e MSR (*Microsoft Research*) [4] para avaliar o banco de dados construído a partir de áudios baixados da internet em conjunto com áudios obtidos de insetos capturas em outro projeto de iniciação científica desenvolvida previamente sob o título: Biometria de insetos: Câmara acústica e aquisição de dados [5]. O modelamento do locutor foi realizado com o auxílio do algoritmo de modelagem GMM-UBM, ao final da execução do código ele nos gera dois gráficos onde podemos ver a confiabilidade do sistema, os gráficos gerados são conforme a Figura 3.

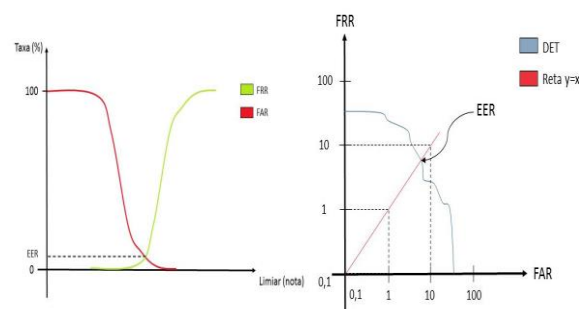


Figura 3 - Curva FRR e FAR / Curva DET.

O EER é uma medida que muitas vezes é utilizada como o principal parâmetro para medir a performance de um sistema biométrico, ela é obtida a partir do cruzamento das curvas FRR (*False Rejection Rate*) e FAR (*False Acceptance Rate*). O segundo gráfico é o DET (*Detection Error Tradeoff*) uma curva de FRR em função de FAR [6].

Na segunda etapa foi utilizado o Raspberry para execução de um novo código, mas os conceitos utilizados são os mesmos usados na primeira etapa com a adição de novas ferramentas para o processamento de áudio como o Audacity que é usado para a remoção do silêncio e aplicação de filtros passa-faixa. O novo front-end pode ser visto na Figura 4.

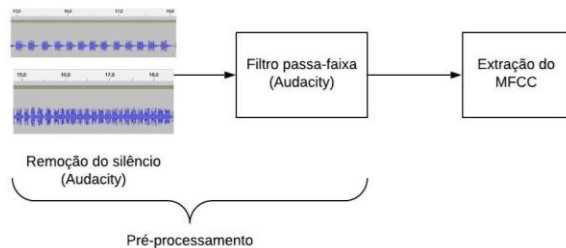


Figura 4 - Front-end no Python.

O *back-end* é baseado na modelagem GMM e comparação utilizando LLR, através da biblioteca *Sklearn* [7]. Após rodar o programa realizado em Python, o código nos retorna como cada locutor foi identificado, portanto, podemos testar a confiabilidade do sistema proposto com a seguinte equação.

$$C = \frac{A}{I} \times 100\% \quad (1)$$

Onde: C é confiabilidade do sistema, A é o número de acertos e I é o total de indivíduos testados.

## 2. Resultados

Os resultados obtidos na primeira etapa com a utilização do MATLAB® são representados na Figura 5.

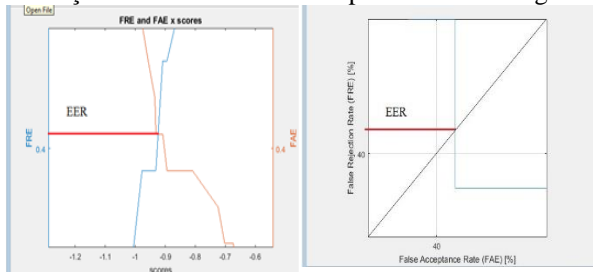


Figura 5 – Resultados em MATLAB®.

Como vemos a taxa de EER está em aproximadamente 45%, portanto, este não é um sistema confiável, isto aconteceu devido à baixa qualidade dos áudios capturados.

O resultado obtido na segunda etapa foi extraído através da execução da fórmula de confiabilidade e a taxa de acerto obtida foi 65%, os erros que tivemos na fase de comparação de resultado foram de espécies que produzem uma frequência muito parecida, além do fato de haver uma variação de frequência produzida pela mesma espécie o que dificulta na hora de gerar um modelo confiável, após todos os testes e com base nos resultados podemos afirmar que é possível a criação de um bom sistema.

## 4. Conclusões

Podemos ver que houve uma melhora em comparação com os resultados obtidos em MATLAB® e Python, isto se deve ao fato de uma melhora na qualidade dos áudios devido as etapas de pré-processamento. Apesar do baixo índice de confiança do sistema, os resultados obtidos foram satisfatórios, pois conseguimos utilizar um sistema que não é muito complexo com o uso de modelamento baseado em GMM e extração de parâmetros MFCC como pilares deste projeto. Estas ferramentas são utilizadas amplamente para identificação de locutor, portanto, com o auxílio das bibliotecas tanto em MATLAB® como em Python faz com que qualquer iniciante entenda a lógica por trás em pouco tempo. Por último, vale ressaltar que o código utilizado é encontrado nas bibliotecas baixadas sofrendo pequenas alterações, isso gera uma gama de possibilidades para aprimoramento do sistema.

Trabalhos futuros serão realizados para aquisição uma maior base de dados, com novas capturas de áudio dos insetos, bem como a utilização de um display no Raspberry Pi® que forneça em tempo real o resultado da identificação de um inseto capturado.

## 5. Referências

- [1] PARADA, Marcelo Gonzaga de Oliveira. **Biometria multimodal baseada nos sinais de voz facial**. Tese (Doutorado em engenharia elétrica) – Centro Universitário FEI, São Bernardo do Campo, 2018.
- [2] PARADA, Marcelo; SANCHES, Ivandro. Comparativo entre GMM-UBM e ivector- PLDA para reconhecimento de locutor na presença de ruído branco, 2016
- [3] REYNOLDS, D. A.; ROSE R. C. **Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models**. *Speech and Audio Processing, IEEE Trans.*, vol. 3(1), p.p. 72-83, 1995.
- [4] SADJADI, S. O.; SLANEY, M.; HECK, L. MSR identity toolbox v1. 0: A MATLAB toolbox for speaker-recognition research, **Speech and Language Processing Technical Committee Newsletter**, vol. 1, no. 4, 2013.
- [5] VICENTE, E. L.; PARADA, M.. BIOMETRIA DE INSETOS VOADORES: Câmara acústica e aquisição de dados, **VII SICFEI**, 2017.
- [6] MARTIN, A. et al. **The DET curve in assessment of detection task performance**. *Proc. Eurospeech 97, Rhodes, Greece, Vol. 4, pp. 1899–1903, Sep. 1997.*
- [7] PEDREGOSA, F. et al. **Scikit-learn: Machine Learning in Python**. *Journal of machine learning research* 12, pp. 2825-2830, Oct. 2011.

## Agradecimentos

Ao Centro Universitário FEI pelo incentivo e apoio ao desenvolvimento acadêmico e pessoal fornecido aos alunos

<sup>1</sup> Aluno de IC do Centro Universitário FEI. Projeto com vigência de 08/18 a 08/19.