

WASIS - Manual do Usuário

01 de Abril de 2018



Conteúdo

Capítulo 1 - Introdução	1
Como Citar.....	1
Componentes do WASIS.....	2
A - Barra de Menu.....	2
B - Barra de Ferramentas.....	2
C - Painel de Bibliotecas.....	2
D - Painel de Visualização do Áudio.....	2
E - Barra de Controles.....	3
Criando uma Biblioteca de Áudio.....	3
Abrindo um Arquivo de Áudio.....	5
Capítulo 2 - Trabalhando com Arquivos Sonoros.....	7
Segmentos de Áudio.....	7
Gravando Arquivos de Áudio.....	8
Funcionalidades da Barra de Controles.....	10
Capítulo 3 - Identificação de Áudios.....	11
Comparação por Força Bruta.....	11
Comparação por Modelos de Classes.....	13

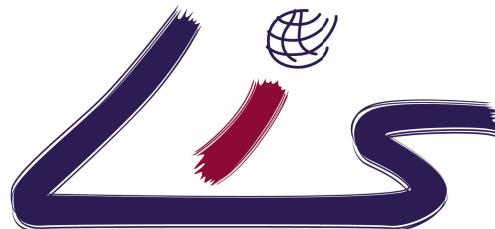
Capítulo 1 - Introdução

WASIS (Wildlife Animal Sound Identification System) é um software amigável, cujo principal objetivo é reconhecer espécies de animais através de seus sons. Em poucas palavras, WASIS extraí informações significativas (descritores) dos sinais sonoros e utiliza uma série de algoritmos de comparação e/ou classificação para comparar registros existentes em diferentes bases de dados a fim de retornar a identificação de uma espécie de animal.

O software WASIS foi desenvolvido na Fonoteca Neotropical Jacques Vielliard (FNJV) em cooperação com o Laboratório de História Natural de Anfíbios Brasileiros (LaHNAB) do Instituto de Biologia, e o Laboratório de Sistemas de Informação (LIS) do Instituto de Computação da Universidade Estadual de Campinas (UNICAMP).



Fonoteca Neotropical Jacques Vielliard



O financiamento deste projeto foi concedido por:

Instituto Virtual de Pesquisas FAPESP - Microsoft Research (Projeto NavScales)
Processo #2011/52070-7 e Processo #2013/02219-0, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)

A versão atual pode ser baixada na página: <http://www.naturalhistory.com.br/wasis.html>

Como citar

Tacioli, L., Medeiros, C. B. & Toledo, L. F. 2018. WASIS: Wildlife Animal Sound Identification System (Version 2.0.0). Fonoteca Neotropical Jacques Vielliard, Unicamp, Brazil. URL: <http://www.naturalhistory.com.br/wasis.html>

Componentes do WASIS

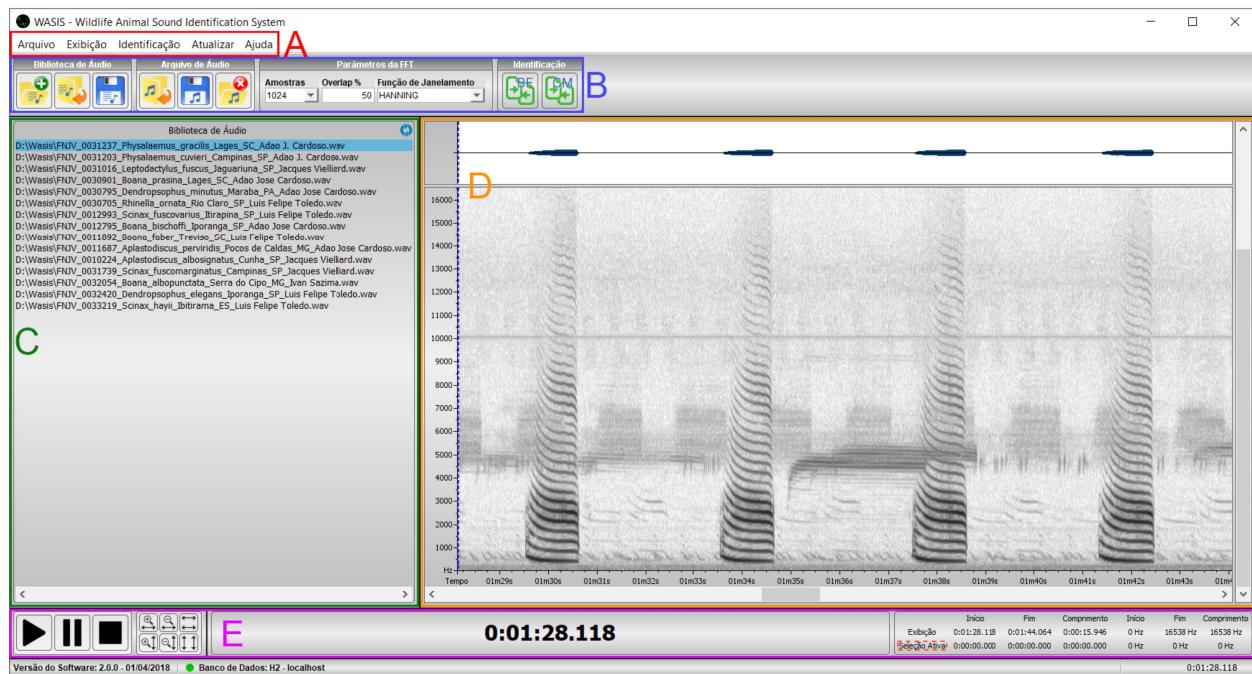


Figura 1 – Tela do WASIS com seus principais componentes destacados

A: Barra de Menu

Localizada no topo da tela, é onde podem ser encontradas todas as funcionalidades para a utilização do WASIS.

B: Barra de Ferramentas

Contém os controles para as operações mais comuns utilizadas no WASIS. A descrição de cada componente é exibida ao posicionar o ponteiro do mouse sobre cada botão.

C: Painel de Bibliotecas

Uma biblioteca é uma lista de arquivos sonoros.

D: Painel de Visualização do Áudio

Quando um arquivo sonoro é carregado, duas representações visuais são exibidas:

- O oscilograma (painel superior) mostra dados de amplitude x tempo;
- O espectrograma (painel inferior) mostra dados do tempo através do eixo horizontal, frequência através do eixo vertical e os níveis de pressão exercida pelo som (Escala Decibel) em um espectro de cores, dos níveis mais fracos (branco) até os níveis mais fortes (preto).

E: Barra de Controles

Contém os seguintes controles:

- Player de áudio;
- Ferramenta de ampliação/redução do oscilograma e espectrograma;
- Tempo atual e frequência visualizada/selecionada.

Criando uma Biblioteca de Áudio

Uma biblioteca é uma lista de arquivos sonoros. É possível criar múltiplas bibliotecas para diferentes análises, mas apenas uma biblioteca pode ser visualizada por vez. Por exemplo, caso desejarmos manter os arquivos sonoros de uma localidade específica, criamos uma nova biblioteca e atribuímos todos os arquivos sonoros a esta biblioteca.

Primeiramente, selecione a opção “Arquivo/Nova Biblioteca de Áudio” na barra de menu ou pressione seu respectivo botão na barra de ferramentas e a seguinte tela será exibida:

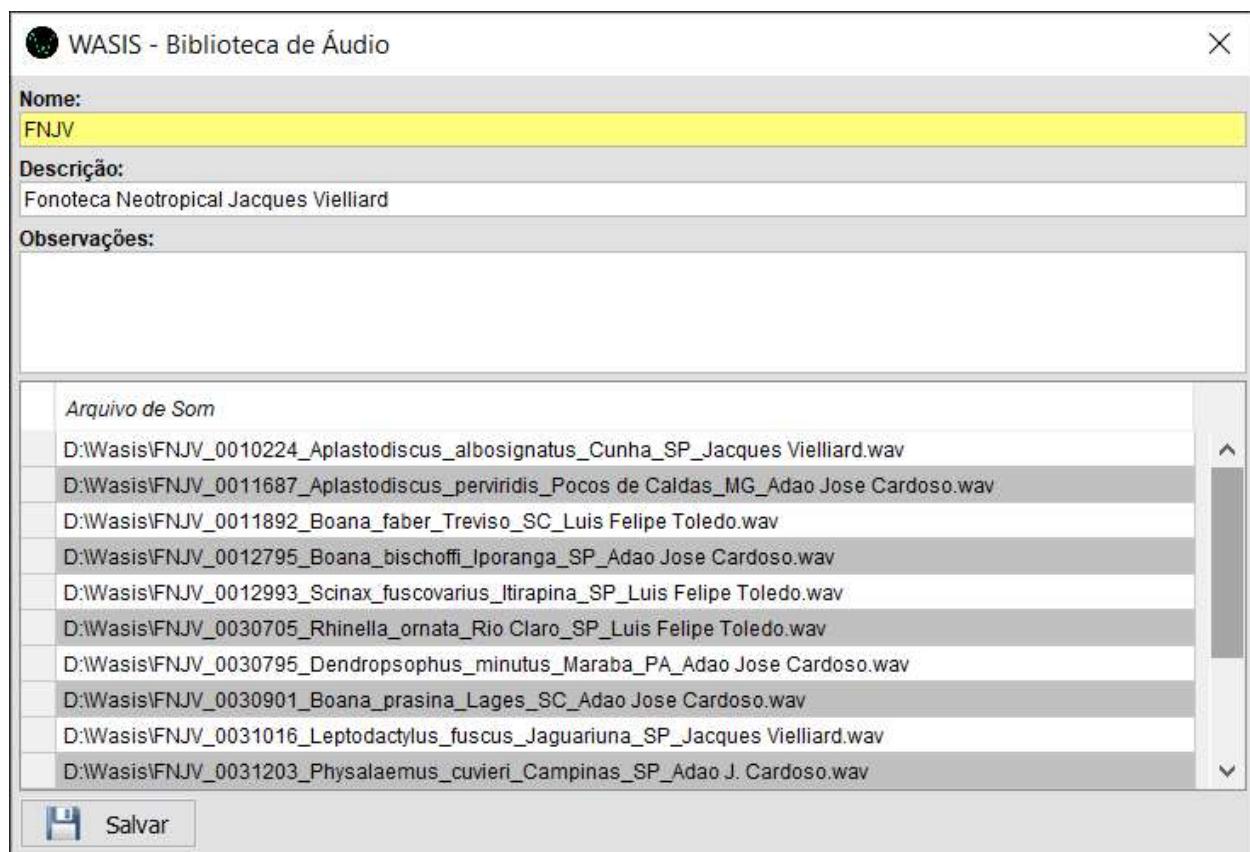


Figura 2 – Tela de Biblioteca com seus respectivos campos e arquivos sonoros já inseridos

Para gravar os dados da nova biblioteca, preencha os seus respectivos campos e pressione o botão “Salvar”.

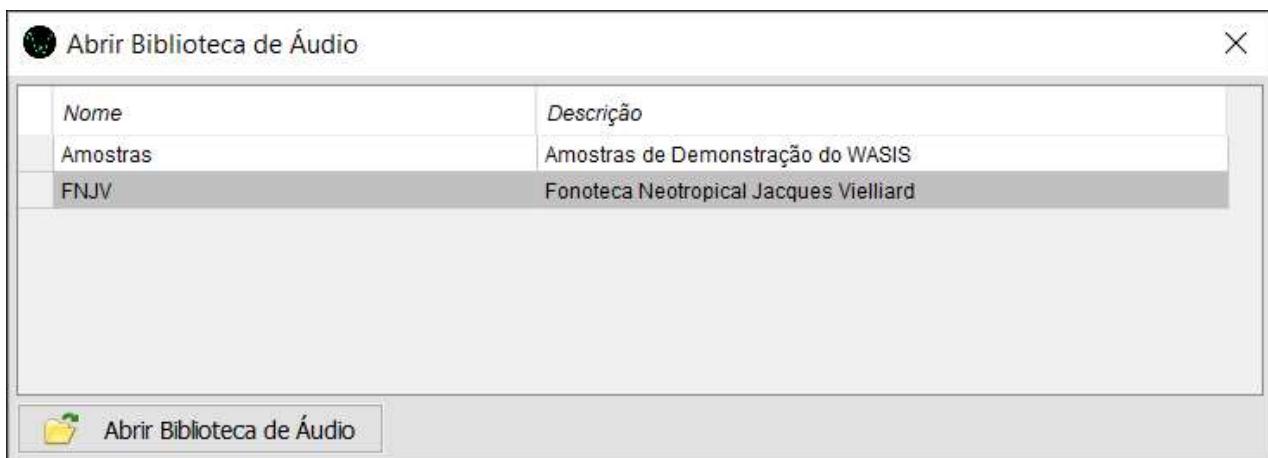


Figura 3 - Lista de bibliotecas de áudio gravadas no banco de dados

Selecione a biblioteca desejada da lista e todos os arquivos sonoros vinculados à biblioteca serão automaticamente carregados no “Painel de Bibliotecas”.



Figura 4 - Painel de Bibliotecas

Notas:

Uma biblioteca não é um recurso obrigatório, mas irá auxiliar na organização dos arquivos sonoros. Um mesmo arquivo sonoro pode ser incluído em diferentes bibliotecas.

Abrindo um Arquivo de Áudio

Para abrir um arquivo sonoro, escolha a opção “Arquivo/Abrir Arquivo de Áudio” ou pressione o botão direito dentro do “Painel de Bibliotecas” e escolha sua respectiva opção.

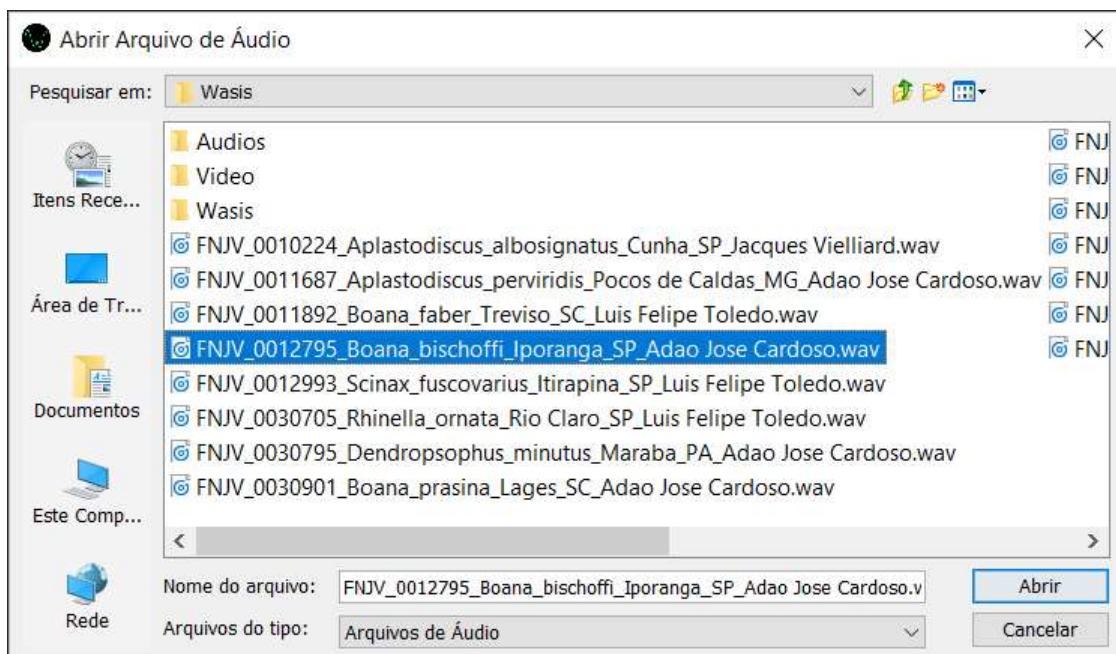


Figura 5 – Tela de abertura de arquivos sonoros

Selecione um arquivo de áudio desejado de uma pasta no computador e o mesmo será carregado automaticamente na biblioteca atual. Ao carregar o arquivo sonoro, o oscilograma e o espectrograma serão exibidos no “Painel de Visualização do Áudio”.

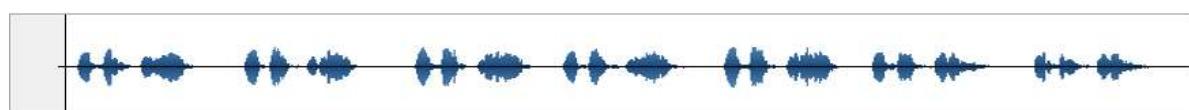


Figura 6 - Oscilograma

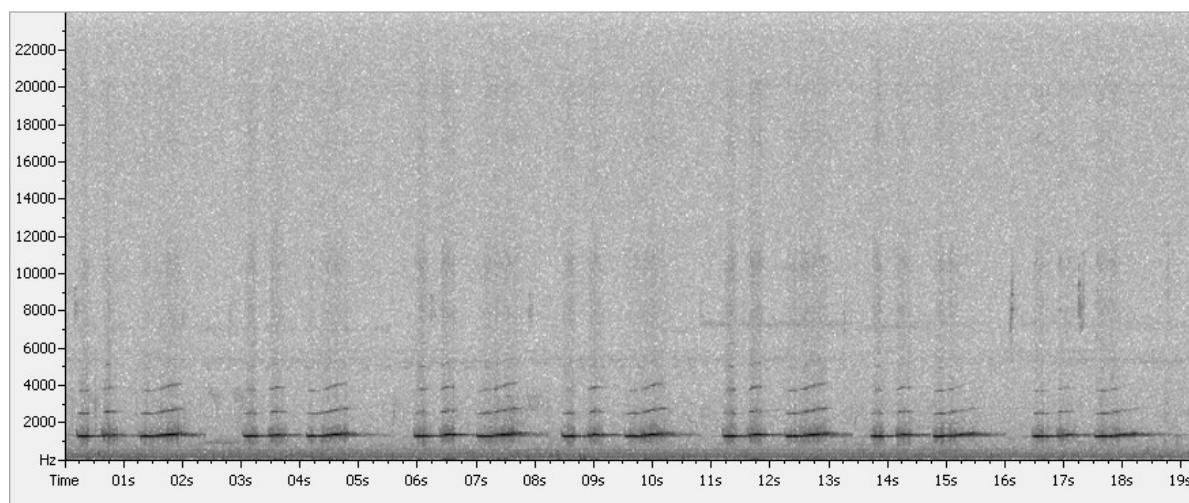


Figura 7 - Espectrograma

O espectrograma é criado baseado no algoritmo FFT (Transformada Rápida de Fourier) e todos seus parâmetros podem ser modificados através dos seus respectivos controles na ‘Barra de Ferramentas’.



Figura 8 - Parâmetros da FFT

Notas:

Por padrão, o WASIS utiliza arquivos sonoros no formato WAV com taxa de amostragem de 44.1 kHz. Em caso de arquivos com diferentes configurações, o sistema irá automaticamente converter o arquivo no formato padrão.

Capítulo 2 – Trabalhando com Arquivos Sonoros

Este capítulo irá explicar como trabalhar com arquivos sonoros em detalhes. As funcionalidades básicas de áudio são controladas através do “Painel de Visualização do Áudio” e a “Barra de Controles”.

O espectrograma representa um papel essencial na análise do som, uma vez que todos os dados extraídos provêm de segmentos selecionados nesta funcionalidade. Para extrair os dados para as análises, selecione a área desejada pressionando o botão esquerdo no mouse no ponto inicial, arrastando até o ponto final e solte o botão do mouse.

Para finalizar a seleção, clique com o botão direito do mouse dentro da área do espectrograma e um menu irá aparecer com as seguintes opções:

- Adicionar novo segmento de áudio: É a identificação da seleção e é rotulada por uma letra e um número. Para cada novo segmento de áudio, a identificação seguirá uma sequência: A1, A2, A3, ..., An;
- Adicionar novo segmento de áudio para uma nova espécie: Se mais de uma espécie é visualizada no espectrograma, esta opção se torna útil para uma melhor organização dos segmentos de áudio.

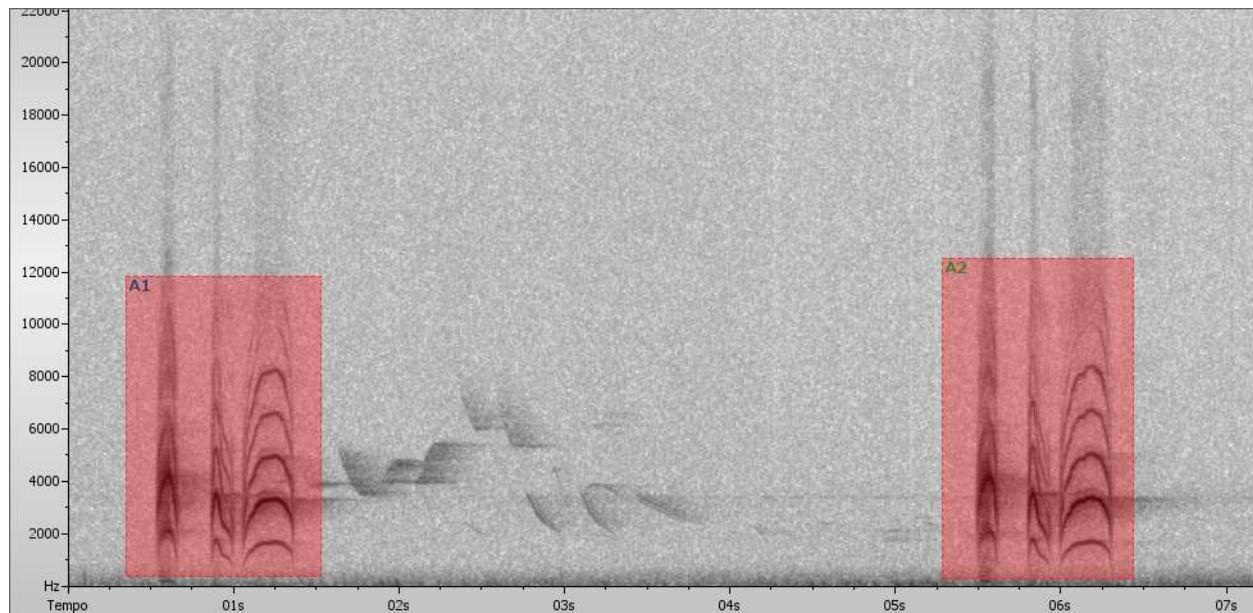


Figura 9 – Segmentos de áudio com suas respectivas identificações.
A1 está assinalado em azul por ser um segmento já gravado no banco de dados,
enquanto A2 está assinalado em verde – não gravado no banco de dados

Há duas opções adicionais quando o botão direito é clicado no espectrograma:

- Aumentar segmento de áudio: Amplifica a área selecionada, mostrando-a em mais detalhes;
- Mostrar lista de segmentos de áudio: Exibe informações de todos os segmentos de áudio do arquivo de áudio, assim como permite a exclusão dos segmentos.

Caminho do Arquivo de Áudio:					
Lista de Segmentos de Áudio:					
	Segmento de Á...	Tempo Inicial	Tempo Final	Frequência Mínima	Frequência Máxima
<input type="checkbox"/>	A1	0:00:00.356	0:00:01.536	377 Hz	11899 Hz
<input type="checkbox"/>	A2	0:00:05.287	0:00:06.456	196 Hz	13181 Hz

 Excluir Segmentos de Áudio

Figura 10 - Lista de segmentos de áudio

Para gravar os segmentos de áudio no banco de dados, escolha a opção “Arquivo/Salvar Arquivo de Áudio” ou seu respectivo botão na barra de ferramentas. Os segmentos serão exibidos e podemos escolher quais deles serão gravados no banco de dados. Marque os segmentos e clique no botão “Salvar Dados do Arquivo de Áudio”.

Caminho do Arquivo de Áudio:					
Segmentos Não Gravados Segmentos Já Gravados					
	Segmento de Áudio	Tempo Inicial	Tempo Final	Frequência Mínima	Frequência Máxima
<input checked="" type="checkbox"/>	A1	0:00:06.003	0:00:07.282	343 Hz	12299 Hz
<input checked="" type="checkbox"/>	A2	0:00:38.327	0:00:39.521	588 Hz	12495 Hz
<input checked="" type="checkbox"/>	A3	0:01:03.268	0:01:05.101	490 Hz	12348 Hz
<input checked="" type="checkbox"/>	A4	0:01:18.608	0:01:20.004	294 Hz	11319 Hz

 Salvar Dados do Arquivo de Áudio

Figura 11 – Tela de escolha dos segmentos de áudio para gravação

Uma nova tela aparecerá com os campos relacionados ao arquivo de áudio. Dados relacionados à classificação científica do animal, bem como informações relacionadas à gravação podem ser preenchidos. Após o preenchimento dos campos, pressione o botão “Salvar Dados”.

Salvar Arquivo de Áudio - Dados X

Informações do Arquivo de Áudio:

Caminho do Arquivo de Áudio:	Número de Voucher:		
ngus sulphuratus\FNJV_0003608_Pitangus_sulphuratus_Itatiaia_RJ_Jacques Vielliard.wav	FNJV 0003608		

Classificação Científica:

Filo:	Classe:	Ordem:
Chordata	Aves	Passeriformes
Família:	Gênero:	Espécie:
Tyrannidae	Pitangus	sulphuratus
Nome Popular:	Nome Popular - Inglês:	
Bem-te-vi	Great Kiskadee	

Informações da Gravação:

País:	Estado:	Cidade:	
Brasil	Rio de Janeiro	Itatiaia	
Dia:	Mês:	Ano:	Hora:
0	0	0	0

Quem Gravou:
Jacques Vielliard

Observações:

Segmento de Áudio Tempo Inicial Frequência Máxima Frequência Mínima Frequência Máxima

A1	0:00:06.003	0:00:07.282	343 Hz	12299 Hz
----	-------------	-------------	--------	----------

< >

Salvar Dados

Figura 12 – Tela de gravação dos dados do arquivo sonoro

Notas:

Muitos dos registros existentes na base de dados do WASIS são criados por seus usuários. Qualquer informação gravada incorretamente, pode gerar resultados incorretos em futuras identificações.

Outras funcionalidades básicas estão localizadas na “Barra de Controles”:

- Player de áudio: Toca todo o arquivo sonoro ou parte dele dependendo dos segmentos de áudio selecionados;



Figura 13 – Comandos do Player de Áudio (Tocar, Pausar and Parar)

- Ferramenta de ampliação/redução: Exibe uma visão ampliada/reduzida do oscilograma e espectrograma;



Figura 14 – Comandos da Ferramenta de ampliação/redução
(Topo: Ampliar - Tempo; Reduzir - Tempo; Resetar - Tempo)
(Baixo: Ampliar - Frequência; Reduzir - Frequência; Resetar - Frequência)

- Tempo atual: Mostra o tempo atual de uma seleção ou quando o arquivo de áudio está sendo tocado.

0:00:18.280

Figura 15 – Tempo atual

- Exibição/Seleção Ativa: Mostra os tempos/frequências iniciais e finais que estão sendo visualizados e selecionados.

	Início	Fim	Comprimento		Início	Fim	Comprimento
Exibição	0:01:11.999	0:01:27.589	0:00:15.590		0 Hz	22050 Hz	22050 Hz
Seleção Ativa	0:01:18.608	0:01:20.004	0:00:01.396		294 Hz	11319 Hz	11025 Hz

Figura 16 – Exibição e Seleção Ativa

Capítulo 3 - Identificação de Áudios

Podemos utilizar duas técnicas para a identificação/classificação de um arquivo sonoro:

Comparação por Força Bruta

Este método de identificação é baseado na comparação direta de informações extraídas dos segmentos de áudio que desejamos identificar com os diversos segmentos já existentes no banco de dados.

Abra um arquivo sonoro e selecione segmentos para a identificação como mostrado anteriormente. Clique na opção “Identificação/Classificação de Áudio - Força Bruta” ou seu respectivo botão na barra de ferramentas e a seguinte tela irá aparecer:

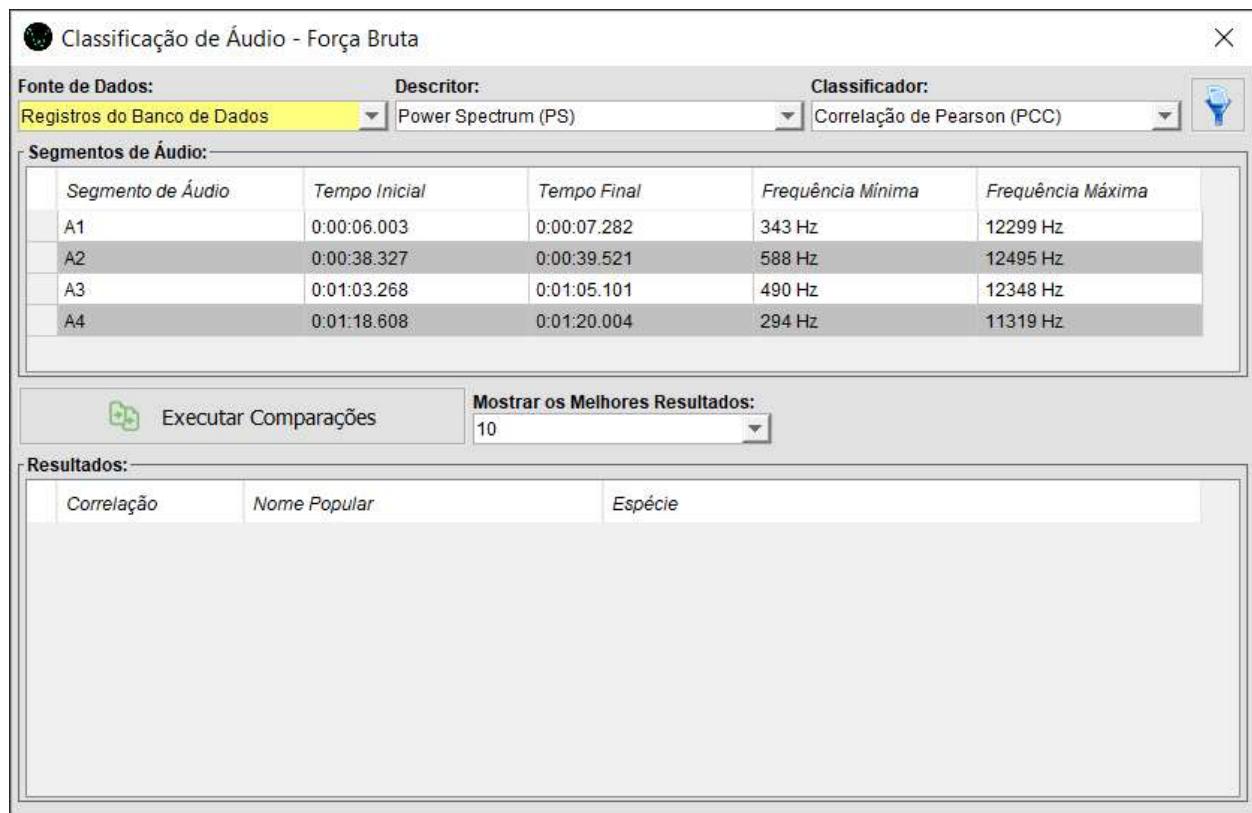


Figura 17 - Classificação de Áudio - Força Bruta

Podemos selecionar diferentes opções de pesquisa que podem variar desde a fonte de dados até os tipos de características extraídas (descritores) dos arquivos sonoros que queremos comparar.

Após a escolha das opções de pesquisa, clique em um dos segmentos de áudio e os resultados da comparação serão exibidos na parte inferior da tela.

Resultados:		
Correlação	Nome Popular	Espécie
0,9418	Bem-te-vi	<i>Pitangus sulphuratus</i>
0,9275	Bem-te-vi	<i>Pitangus sulphuratus</i>
0,9187	Bem-te-vi-rajado	<i>Myiodynastes maculatus</i>
0,9098	Bem-te-vi-rajado	<i>Myiodynastes maculatus</i>
0,9084	Bem-te-vi	<i>Pitangus sulphuratus</i>
0,9072	Bem-te-vi-rajado	<i>Myiodynastes maculatus</i>
0,9054	Tico-tico	<i>Zonotrichia capensis</i>
0,8938	Gavião-carijó	<i>Rupornis magnirostris</i>
0,8929	Enferruiado	<i>I. lathrotriccus euleri</i>

Figura 18 - Resultados da Comparaçāo por Força Bruta

No exemplo acima, os resultados da comparação são baseados nas análises de distribuição de potência do som entre diversas faixas de frequência (Power Spectrum). A Correlação de Pearson (PCC) calcula as similaridades entre os registros e uma classificação final é apresentada – quanto maior a correlação, maior a probabilidade de ser a espécie identificada corretamente.

Um duplo clique nos resultados da comparação exibe um gráfico comparando os dados entre o segmento a ser identificado e o existente na base dados.

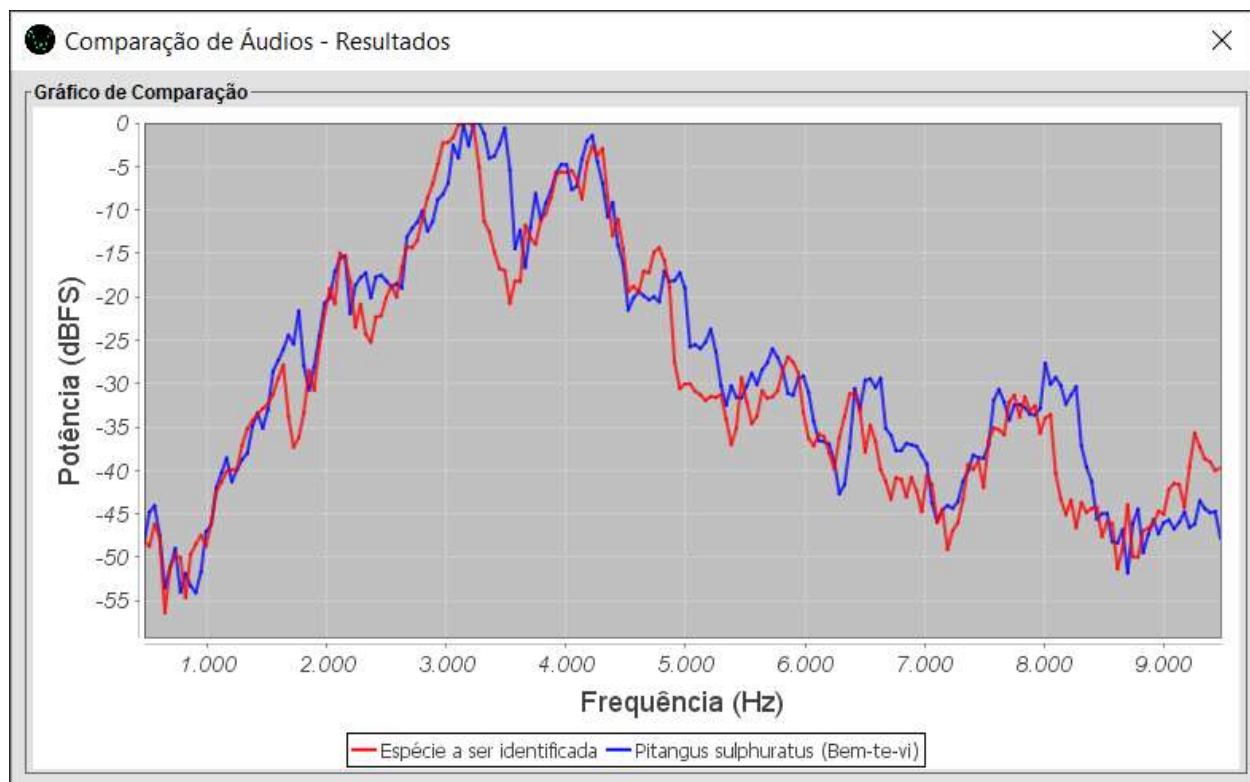


Figura 19 – Arquivos sonoros de uma mesma espécie apresentam curvas de dados com padrões semelhantes

Comparação por Modelos de Classes

Geralmente, este método utiliza algoritmos de aprendizado de máquina supervisionado para a identificação de sons. O propósito principal deste tipo de algoritmo é projetar um método de classificação que possa melhor predizer as classes/rótulos para dados desconhecidos (um arquivo de áudio que queremos identificar). Este processo é similar à habilidade do cérebro humano em diferenciar entre uma ampla variedade de sons e atribuí-los a sons ouvidos anteriormente. Estes algoritmos processam conjuntos de dados de diversas espécies ao mesmo tempo e cria um modelo final para cada espécie.

Clique na opção “Identificação/Classificação de Áudio - Modelos de Classes” ou seu respectivo botão na barra de ferramentas e a seguinte tela irá aparecer:

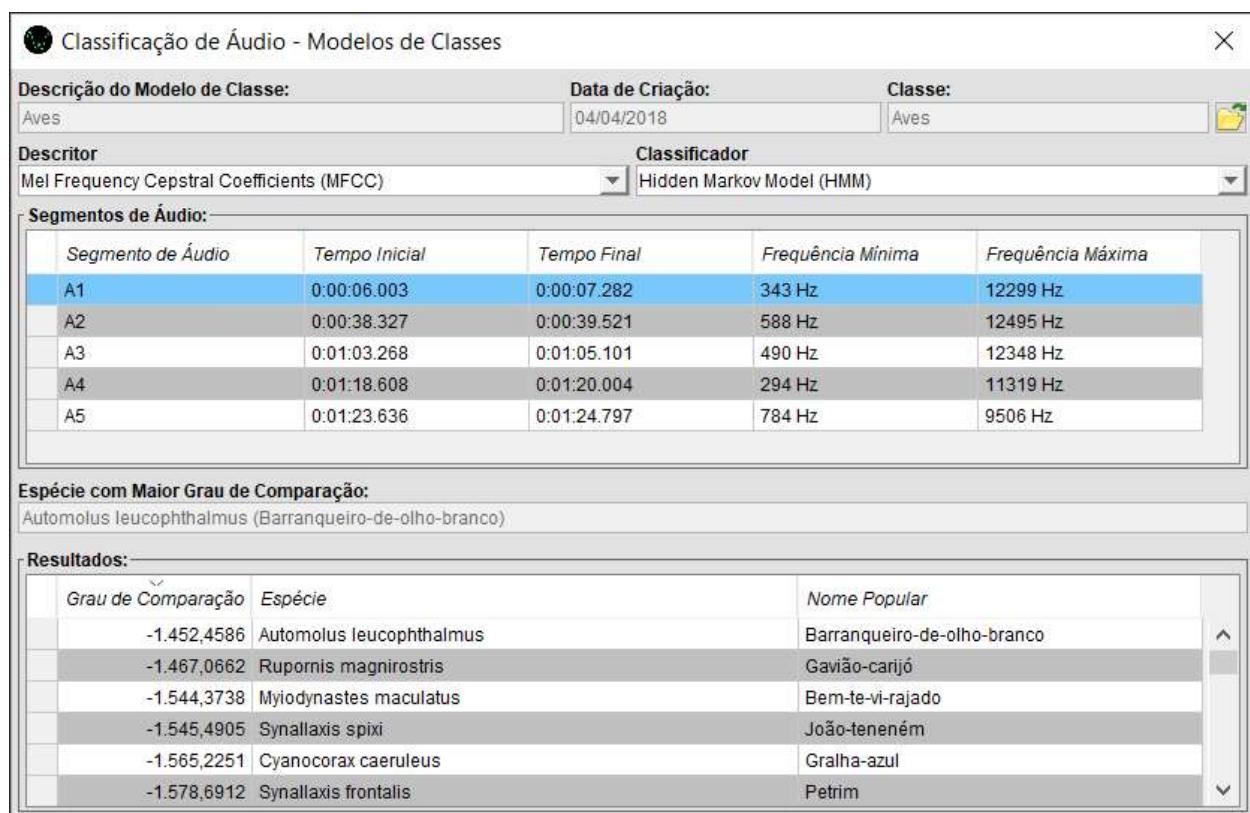


Figura 20 - Classificação de Áudio - Modelos de Classes

Escolha um grupo de classes já treinado anteriormente juntamente com as demais opções de classificação, selecione um segmento de áudio para a identificação e os resultados da identificação serão exibidos na parte inferior da tela.