

Comparação de Métodos de Classificação de Ruído Acústico

Antônio Nascimento, Felipe Farias e Marília Alves

Instituto Militar de Engenharia

Agosto de 2017

- 1 Introdução
 - Processamento de Sinais Acústicos em Ruído
 - Objetivos
- 2 Técnicas Utilizadas
 - Validação Cruzada
 - Extração de Atributos do Áudio
- 3 Experimentos
 - Condições Iniciais
 - Métodos Utilizados
- 4 Considerações Finais

Ruído em Tarefas de PDS

- Tarefas de Processamento de Sinais Acústicos
 - Reconhecimento de Locutor
 - Localização
 - Reconhecimento de Fala
- Ruído é o **principal** desafio.
 - Queda de rendimento.

Estudar e Analisar Ruído

- Detecção (presença/ausência)
- Classificação
 - Inclusão em Ferramentas de Realce

Objetivos do Trabalho

- Comparar desempenho de técnicas de classificação de áudio conhecidas na literatura na tarefa de classificação de ruído.

Metodologia

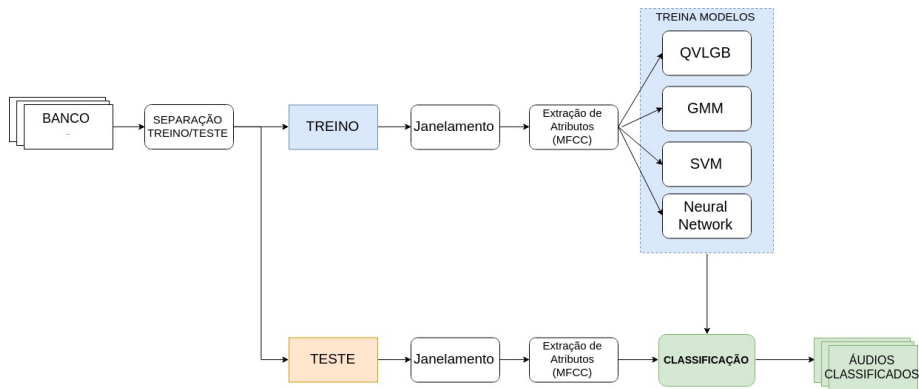


Figura: Diagrama da Classificação

A base de dados NOISEX-92

- 15 classes
- Formato .wav
- Tamanho: 3m55s

Tabela: Classes na base de dados NOISEX.

Babble	Buccaneer 1	Buccaneer 2	Destroyer Engine	Destroyer Operations Room
F16	Factory Floor 1	Factory Floor 2	HF Channel	Leopard
M109	Machine Gun	Pink Noise	Volvo	White Noise

Validação Cruzada



Figura: Diagrama da Classificação

Mel-Frequency Cepstral Coefficient (MFCC)

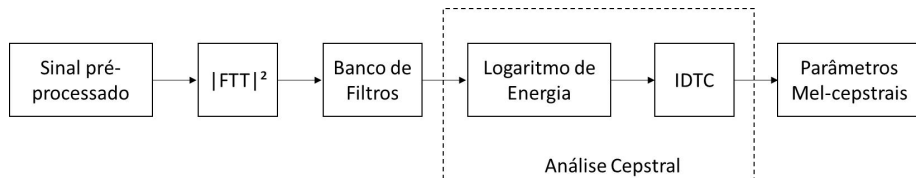


Figura: Extração do MFCC

Condições Iniciais

- Entrada dos classificadores:
 - 15 classes;
 - 275370 amostras;
 - 39 coeficientes;
- Ferramentas Utilizadas:
 - MATLAB;
 - toolbox Neural Network;
 - toolbox SVM;
 - Voicebox;

QV LBG (intro)

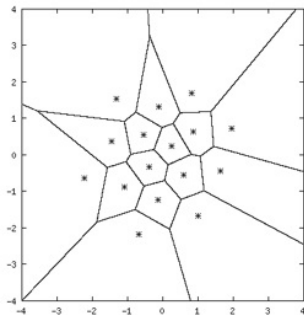


Figura: Exemplo de QV de 2 dimensões.

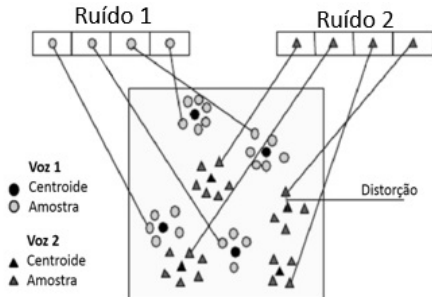


Figura: Diagrama que ilustra um processo QV.

QV LBG (resultados)

Tabela: Matriz de confusão do classificador QV LBG

QV LBG	Babble	Buccaneer1	Buccaneer2	Destroyerengine	Destroyerops	F16	Factory1	Factory2	Hfchannel	Leopard	M109	Machinegun	Pink noise	Volvo	White Noise
Babble	86,011	0,007	0,000	0,799	0,530	0,225	5,244	4,808	0,000	0,182	1,881	0,305	0,007	0,000	0,000
Buccaneer1	0,000	97,109	0,000	0,000	0,000	0,567	2,012	0,000	0,000	0,000	0,007	0,000	0,305	0,000	0,000
Buccaneer2	0,000	0,029	98,860	0,000	0,015	0,022	0,000	0,000	0,000	0,000	0,000	0,000	0,399	0,000	0,675
Destroyerengine	0,029	0,000	0,000	99,644	0,000	0,000	0,327	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Destroyerops	0,007	0,298	0,429	0,000	91,793	0,022	2,005	0,000	0,000	0,000	2,796	0,007	2,644	0,000	0,000
F16	86,018	0,007	0,000	0,799	0,530	0,225	5,244	4,808	0,000	0,182	1,881	0,305	0,007	0,000	0,000
Factory1	4,670	5,295	1,344	2,491	2,716	0,966	59,057	9,675	0,000	0,000	0,218	0,000	13,466	0,000	0,102
Factory2	4,416	0,000	0,000	0,015	0,000	0,610	0,770	94,095	0,000	0,000	0,007	0,007	0,000	0,080	0,000
Hfchannel	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	100,000	0,000	0,000	0,000	0,000	0,000	0,000
Leopard	0,153	0,000	0,000	0,000	0,000	0,000	0,000	0,044	0,000	98,693	0,000	0,174	0,000	0,937	0,000
M109	0,022	0,000	0,000	0,000	3,167	0,000	0,436	0,116	0,000	0,029	93,892	2,310	0,022	0,007	0,000
Machinegun	0,000	0,015	0,000	0,000	3,850	0,000	0,044	0,000	0,000	43,921	0,857	7,452	38,655	5,208	0,000
Pink noise	0,000	0,036	0,094	0,000	0,109	0,000	0,000	0,000	0,000	0,000	0,000	0,000	99,760	0,000	0,000
Volvo	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,065	0,000	0,182	0,007	8,963	0,000	90,783	0,000
White Noise	0,000	0,000	0,051	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	99,949

Gaussian Mixture Model (intro)

$$p(\vec{x}|\lambda) = \sum_{k=1}^K p_k b_k(\vec{x}) \quad (1)$$

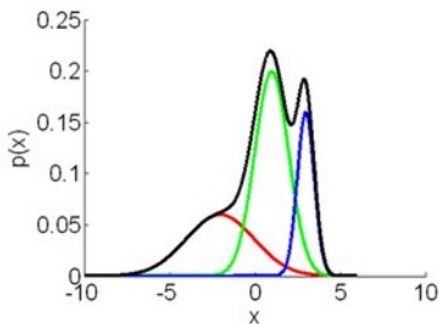


Figura: Ilustração do funcionamento do GMM.

Gaussian Mixture Model (resultados)

Tabela: Matriz de confusão do classificador GMM

GMM	Babble	Buccaneer1	Buccaneer2	Destroyerengine	Destroyerops	F16	Factory1	Factory2	Hfchannel	Leopard	M109	Machinangun	Pink noise	Volvo	White Noise
Babble	98,860	0,000	0,000	0,015	0,087	0,000	0,559	0,436	0,000	0,000	0,036	0,000	0,000	0,007	0,000
Buccaneer1	0,000	99,266	0,000	0,000	0,007	0,116	0,574	0,000	0,000	0,000	0,000	0,000	0,036	0,000	0,000
Buccaneer2	0,007	0,000	99,688	0,000	0,087	0,000	0,102	0,000	0,000	0,000	0,000	0,000	0,022	0,000	0,094
Destroyerengine	0,073	0,000	0,000	99,746	0,000	0,000	0,182	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Destroyerops	0,065	0,022	0,029	0,000	97,785	0,007	1,547	0,000	0,000	0,000	0,356	0,000	0,189	0,000	0,000
F16	0,000	0,182	0,000	0,000	0,015	98,932	0,843	0,022	0,000	0,000	0,007	0,000	0,000	0,000	0,000
Factory1	0,588	0,501	0,044	0,189	0,763	0,552	91,182	4,467	0,000	0,000	0,087	0,000	1,627	0,000	0,000
Factory2	0,646	0,000	0,000	0,000	0,007	0,051	4,278	94,988	0,000	0,007	0,000	0,000	0,000	0,022	0,000
Hfchannel	0,007	0,000	0,000	0,000	0,000	0,000	0,000	0,000	99,993	0,000	0,000	0,000	0,000	0,000	0,000
Leopard	0,116	0,000	0,000	0,000	0,000	0,000	0,000	0,007	0,000	99,005	0,000	0,537	0,000	0,334	0,000
M109	0,015	0,000	0,000	0,000	0,632	0,000	0,436	0,000	0,000	0,007	98,903	0,007	0,000	0,000	0,000
Machinangun	0,007	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,203	0,007	99,680	0,000	0,102	0,000
Pink noise	0,000	0,036	0,073	0,000	0,153	0,000	1,286	0,000	0,000	0,000	0,007	0,000	98,446	0,000	0,000
Volvo	0,015	0,000	0,000	0,000	0,000	0,000	0,000	0,087	0,000	0,102	0,007	0,036	0,000	99,753	0,000
White Noise	0,000	0,000	0,044	0,000	0,000	0,000	0,007	0,000	0,000	0,000	0,000	0,000	0,000	0,000	99,949

Support Vector Machines (intro)

- texto aqui

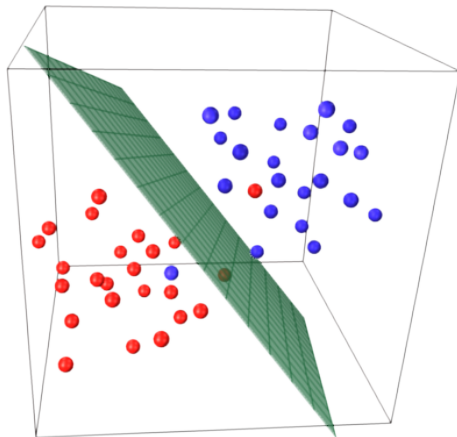


Figura: Ilustração do funcionamento de um SVM.

Support Vector Machines (resultados)

Rede Neural (intro)

- texto aqui



Figura: Esquema de funcionamento de uma rede neural.

Rede Neural (resultados)

Comparação

Tabela: Acurácia dos métodos estudados por classe.

Class	VQ-LBG	GMM	NN	SVM
Babble	88,2%	98,9%	98,7%	99,4%
Bucanneer 1	96,6%	99,1%	99,2%	99,7%
Bucanneer 2	98,8%	99,7%	99,9%	99,9%
Destroyer Engine	99,8%	99,7%	99,8%	99,9%
Destroyer Ops	90,8%	96,9%	98,7%	99,5%
F16	95,2%	99,1%	99,5%	99,8%
Factory 1	59,3%	87,6%	93,5%	93,8%
Factory 2	93,9%	95,0%	95,0%	94,7%
HF Channel	100,0%	100,0%	100,0%	100,0%
Leopard	99,1%	99,6%	99,6%	99,7%
M109	94,1%	99,3%	99,2%	99,6%
Machine Gun	7,1%	99,5%	99,7%	99,9%
Pink Noise	99,7%	98,0%	97,8%	97,8%
Volvo	90,8%	99,4%	99,7%	99,9%
White Noise	99,9%	99,9%	100,0%	100,0%
OVERALL	89,2%	98,0%	99,5%	99,7%

Conclusões

- Três dos quatro métodos tiveram desempenho acima de 95%.
 - Melhor: SVM.
 - Pior: QV LBG.
- Pontos Negativos:
 - Alta Complexidade Computacional.
 - Necessidade de grande base de dados.

Trabalhos Futuros

- Estender a comparação a outros métodos de classificação;
- Aumentar a diversidade da base de dados;
- Investigar desempenho de comitê de classificação;
- Investigar técnicas de Deep Learning.

Obrigado!

- Obrigado pela atenção!