

Identificação de Locutor

Ramon Duarte de Melo & André Ribeiro Queiroz

Universidade Federal do Rio de Janeiro

ramonduarte@poli.ufrj.br & handre_queiroz@poli.ufrj.br

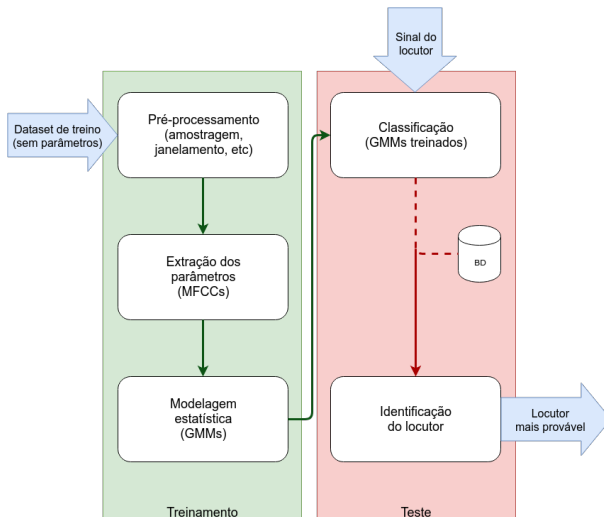
Código e apresentação disponíveis em: <https://github.com/ramonduarte/procvoz>

9 de julho de 2019

Sumário

- 1 Arquitetura
- 2 Dados de entrada
- 3 Pré-processamento
- 4 Treinamento
- 5 Resultados

Visão Geral do Sistema



Sinais de entrada (treino e teste)

- Áudios de entrada amostrados em 16kHz
- 16 bits por amostra
- Duração variável entre 10 e 60 segundos
- Dataset de treinamento: 34 locutores, 5 frases para cada locutor, 20-30 segundos
- Dataset de teste: os mesmos 34 locutores, 5 frase para cada, ≈ 10 segundos

Amostragem e Janelamento

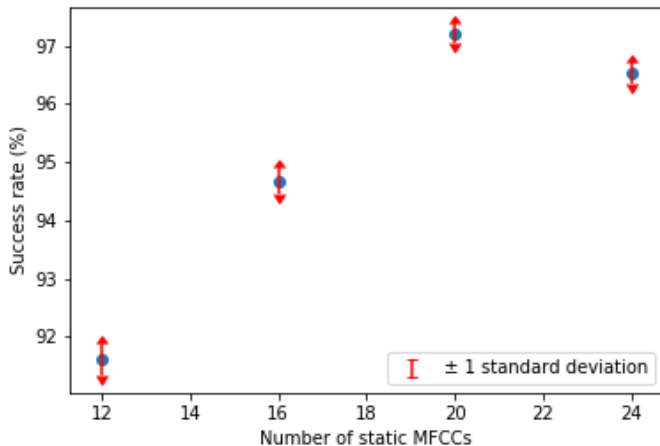
- 1 O sinal de áudio foi dividido em quadros que variaram de 20 a 30 ms para encontrar estacionariedade.
- 2 Foram testados dois tipos similares de janelamento: a função de Hamming e a função de Blackman.
- 3 Em cima disto, foram testadas também as seguintes distâncias entre janelas: 5 ms, 10 ms, 15 ms e 20 ms.
- 4 Um total de 96 combinações foram executadas.

Parametrização

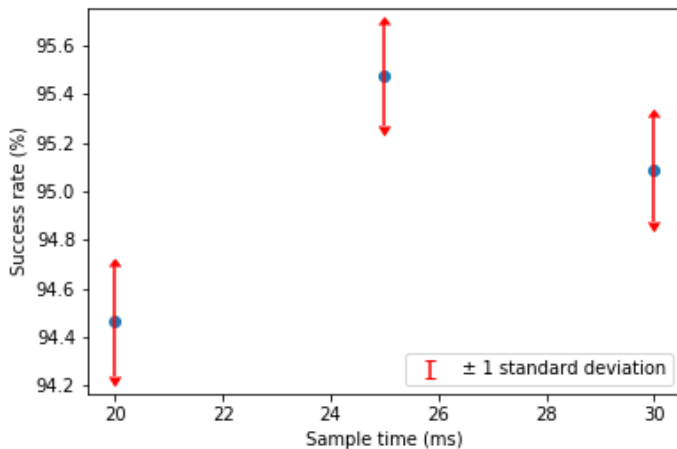
- O programa aceita um número arbitrário de MFCCs, desde que o número de coeficientes estáticos seja igual ao de coeficientes dinâmicos.
- Foram testados $12 + 12$, $16 + 16$, $20 + 20$ e $24 + 24$ coeficientes.

	mfccs	window	dist	fn	speaker	guess	correct guess
1188	12	0.025	0.015	Hamming	belmontguy	belmontguy	True
1189	12	0.025	0.015	Hamming	belmontguy	belmontguy	True
1190	12	0.025	0.020	Hamming	anthonyshaller	arvala	False
1191	12	0.025	0.020	Hamming	anthonyshaller	arvala	False
1192	12	0.025	0.020	Hamming	anthonyshaller	anthonyshaller	True

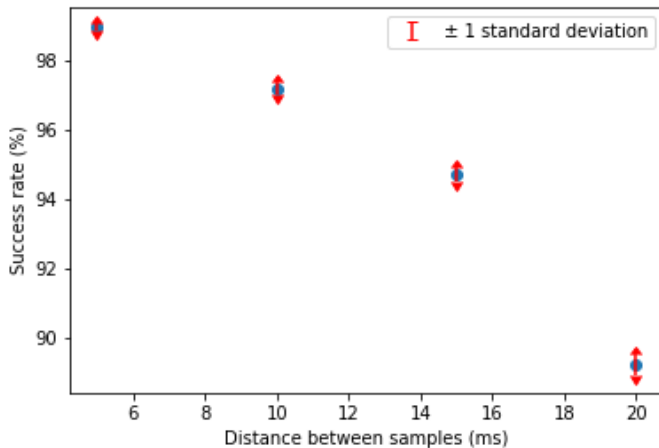
Análise dos parâmetros: número de MFCCs



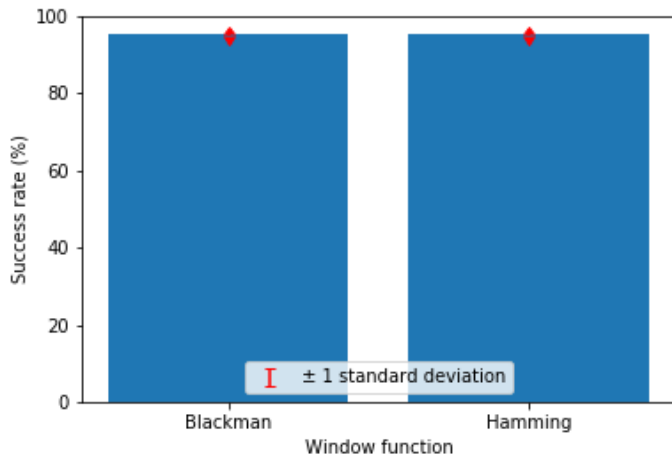
Análise dos parâmetros: intervalo de amostragem



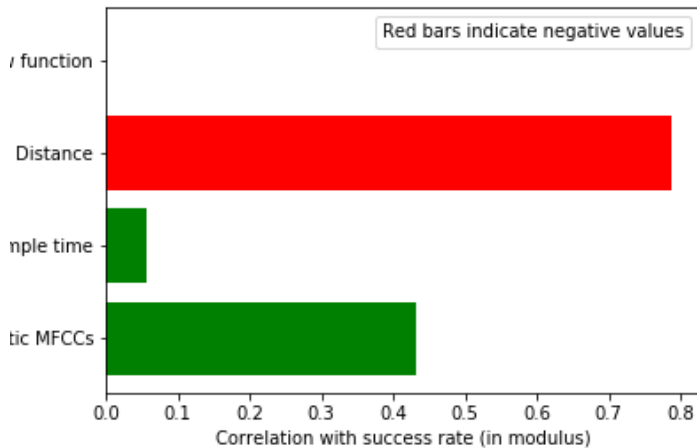
Análise dos parâmetros: distância entre janelas



Análise dos parâmetros: função de janelamento



Análise dos parâmetros: correlação linear



Análise dos parâmetros: piores e melhores combinações

	mfccs	window	dist	fn	count
6	12	0.020	0.020	Blackman	0.823529
7	12	0.020	0.020	Hamming	0.823529
14	12	0.025	0.020	Blackman	0.847059
15	12	0.025	0.020	Hamming	0.847059
22	12	0.030	0.020	Blackman	0.847059
23	12	0.030	0.020	Hamming	0.847059
30	16	0.020	0.020	Blackman	0.852941
31	16	0.020	0.020	Hamming	0.852941
38	16	0.025	0.020	Blackman	0.882353
39	16	0.025	0.020	Hamming	0.888235
56	20	0.025	0.005	Blackman	1.000000
57	20	0.025	0.005	Hamming	1.000000
88	24	0.030	0.005	Blackman	1.000000
89	24	0.030	0.005	Hamming	1.000000