

Referências Bibliográficas

- [1] ALCAIM, A.. **Processamento de Voz e Imagem**. PUC Rio, 1999.
- [2] ATAL, B.. **Automatic Recognition of speakers from Their Voices**. Proceeding of the IEEE, Vol. 64, No. 4, pp. 460-475, Apr 1976.
- [3] BESACIER, L.; BONASTRE, J.. **Subband Architecture for Automatic Speaker Recognition** . Proceeding of Signal Processing, pp.1245-1259, 2000.
- [4] BEZERRA, M. D. R.. **Reconhecimento automatico de locutor para fins forenses, utilizando técnicas de redes neuronais**. Master's thesis, Instituto Militar De Engenharia, 1994.
- [5] CAMPBELL, J. P.. **Speaker Recognition: A Tutorial** . Proceeding of the IEEE, Vol. 85, No. 9, pp. 1437-1462, Sept 1997.
- [6] D'ALMEIDA, F. Q.. **Técnicas Eficientes de Identificação Automática de Locutores**. PhD thesis, Universidade de Brasília, 2009.
- [7] DELLER, J.; PROAKIS, J. ; HANSEN, J.. **Discrete-Time Processing of Speech Signals**. IEEE Press, 2000.
- [8] DIMITRIADIS, D.; MARAGOS, P. ; POTAMIOS, A.. **Auditory Teager Energy Cepstrum Coefficients for Robust Speech Recognition**. Proceeding of European Speech Processing Conference-Interspeech, pp.3013-3016, Sept 2005.
- [9] GAJIC, B.. **Feature extraction for Automatic Speech Recognition in Noisy Acoustic Environment**. PhD thesis, Norwegian University of Science and Technology, 2002.
- [10] GAJIC, B.; PALIWAL, K.. **Robust Speech Recognition in Noisy Environments Based on Subband Spectral Centroid Histograms**. IEEE Transactions on Speech and Audio Processing, pp. 1-9, 2006.
- [11] GHITZA, O.. **Auditory Models and Human Performance in Tasks Related to Speech Coding and Speech Recognition** . IEEE

- Transactions on Speech and Audio Processing, Vol. 2, No. 1, pp. 115-132, Jan 1994.
- [12] HIGGINS, J. E. AND DAMPER, R. I. D. T. J.. **Information Fusion for Subband-HMM Speaker Recognition**. Proceeding of IEEE Int. Joint Conf. on Neural Networks, Vol. 2, pp. 1504-1509, Jul 2001.
- [13] IKBAL, S.; MISRA, H. ; BOURLARD, H.. **Phase Autocorrelation (PAC) derived robust speech features**. Proceeding of IEEE Int. conf. on Acoust. Speech, and Signal Processing (ICASSP), pp. II133-II136, April 2003.
- [14] IRINO, T.; PATTERSON, R. D.. **A time-domain level-dependent auditory filter: The gammachirp**. Proceeding of Acoustical Society of America, Vol. 101, pp. 412-419, Jan 1997.
- [15] JAYANT, M. N.. **Speaker Verification: A Tutorial**. IEEE Communications Magazine, pp. 42-47, Jan 1990.
- [16] **Noisex-92**. <http://www.speech.cs.cmu.edu/comp.speech/noisex.html>.
- [17] PICONE, J. W.. **Signal Modeling Techniques in Speech Recognition**. Proceeding of the IEEE, Vol. 81, pp. 1215-1247, 1993.
- [18] RABINER, L.; SHAFER, R. W.. **Digital Processing of Speech Signals**. Prentice Hall, 1978.
- [19] REYNOLDS, D. A.. **Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models**. IEEE Transactions on Speech and Audio Processing, Vol. 3, No. 1, pp. 72-83, Jan 1995.
- [20] RIQUELME CUADROS, C.. **Reconhecimento de voz e de locutor em ambientes ruidosos: comparação das técnicas zcpa e mfcc**. Master's thesis, Universidade Federal Fluminense, 2007.
- [21] SAKKA, Z. AND KACHOURI, A.; MEZGHANI, A. AND SAMET, M.. **A New Method for Speech Denoising and Speaker Verification Using Subband Architecture**. . First International Symposium on Control, Communications and Signal Processing, pp. 37-40, Mar 2004.
- [22] STEVENS, S. S. AND VOLKMAN, J.. **The Relation of Pitch to Frequency: A Revised Scale**. . Journal of Psychology, Vol. 53, pp. 329, Jul 1940.

- [23] VALE, E.; CUNHA, A. ; ALCAIM, A.. **Robust Text-Independent Speaker Identification Using Multiple Subband-Classifiers in Colored Noise Environment** . Proceeding of 15th Int.conf.on Systems,Signal and Image Processing(IWSSIP), pp.275-278, 2008.
- [24] YOUNG, S.; EVERMANN, G. ; GALES, M.. **The HTK Book (for HTK Version 3.3)**. 2005.

A

Resultados Experimentais Detalhados

Tabela A.1: Comparação dos porcentagens de acertos obtidas pelos atributos utilizados para o reconhecimento automático del locutor independente do texto, a avaliação foi feita com a base de dados TIMIT para locuções limpas, e com ruído de fábrica para diferente relação sinal ruído SNR.

Técnica utilizada	Porcentagem de Acerto[%]					
	Limp	Fábrica				
		SNR [dB]				
		20	15	10	5	0
MFCC	92.46	81.83	62.22	31.27	6.11	0.87
SSCH	91.59	77.94	59.60	30.87	8.81	1.43
TECC	89.68	71.27	52.78	26.03	6.43	1.27
PAC-MFCC	81.83	72.70	61.43	43.81	20.16	5.24
2 Sub-bandas ^a	85.55	74.60	64.40	44.76	21.66	5.32
4 Sub-bandas ^b	79.68	64.68	50.87	27.70	9.20	2.69

^aPAC-MFCC com Classificador em Sub-Bandas utilizando 2 sub-bandas

^bPAC-MFCC com Classificador em Sub-Bandas utilizando 4 sub-bandas

Tabela A.2: Comparação dos porcentagens de acertos obtidas pelos atributos utilizados para o reconhecimento automático del locutor independente do texto, a avaliação foi feita com a base de dados TIMIT para locuções limpas, e com ruído falatório para diferente relação sinal ruído SNR.

Técnica utilizada	Porcentagem de Acerto[%]					
	Limp	Falatório				
		SNR [dB]				
		20	15	10	5	0
MFCC	92.46	84.44	72.70	44.29	13.10	2.30
SSCH	91.59	80.56	68.81	38.97	9.52	1.27
TECC	89.68	72.30	52.54	25.40	7.46	1.19
PAC-MFCC	81.83	73.49	63.17	45.56	24.92	8.33
2 Sub-bandas ^a	85.55	73.80	63.33	45.39	23.41	6.98
4 Sub-bandas ^b	79.68	61.43	51.19	31.34	11.42	2.62

^aPAC-MFCC com Classificador em Sub-Bandas utilizando 2 sub-bandas

^bPAC-MFCC com Classificador em Sub-Bandas utilizando 4 sub-bandas

Tabela A.3: Comparação dos porcentagens de acertos obtidas pelos atributos utilizados para o reconhecimento automático del locutor independente do texto, a avaliação foi feita com a base de dados TIMIT para locuções limpas, e com ruído branco para diferente relação sinal ruído SNR.

Técnica utilizada	Porcentagem de Acerto[%]					
	Limpo	Branco				
		SNR [dB]				
		20	15	10	5	0
MFCC	92.46	70.32	42.46	15.95	4.21	0.63
SSCH	91.59	69.29	46.03	20.56	6.38	1.27
TECC	89.68	70.32	50.24	28.41	8.89	1.9
PAC-MFCC	81.83	67.38	49.52	24.13	6.43	1.43
2 Sub-bandas ^a	85.55	73.33	62.22	44.92	23.25	9.84
4 Sub-bandas ^b	79.68	62.06	48.1	29.68	20.31	6.35

^aPAC-MFCC com Classificador em Sub-Bandas utilizando 2 sub-bandas

^bPAC-MFCC com Classificador em Sub-Bandas utilizando 4 sub-bandas

Tabela A.4: Comparação dos porcentagens de acertos obtidas pelos atributos utilizados para o reconhecimento automático de locutor independente do texto, a avaliação foi feita com a base de dados TIMIT para locuções limpas, e com ruído de pink para diferente relação sinal ruído SNR.

Técnica utilizada	Porcentagem de Acerto[%]					
	Limpo	Pink				
		SNR [dB]				
		20	15	10	5	0
MFCC	92.46	79.05	55.71	22.22	4.44	0.63
SSCH	91.59	74.52	52.62	27.38	7.70	0.95
TECC	89.68	69.84	49.84	23.17	5.79	0.16
PAC-MFCC	81.83	71.90	61.35	42.06	17.78	5.63
2 Sub-bandas ^a	85.55	74.44	66.35	46.66	23.76	6.19
4 Sub-bandas ^b	79.68	64.36	51.03	29.76	12.22	3.09

^aPAC-MFCC com Classificador em Sub-Bandas utilizando 2 sub-bandas

^bPAC-MFCC com Classificador em Sub-Bandas utilizando 4 sub-bandas

B

Tempo de Simulação por Atributo

A seguinte tabela apresenta o tempo de simulação para cada um dos atributos utilizados. Para esta tarefa foi utilizado um computador com processador Intel Core I5 de 2.66Ghz, memória Ram de 8Ghz e Sistema Operativo Windows 7. Além disso foram utilizados os programas Matlab 7.8.0(R2009a) e Htk 3.4.

Tabela B.1: Comparação dos tempos de simulação obtidos pelos atributos utilizados para o reconhecimento automático del locutor independente do texto.

Técnica utilizada	Tempo (min)			Tempo Total (horas)
	Extração dos Atributos		Reconhecimento	
	Treinamento	Teste		
MFCC	1.89	0.45	125	2.12
SSCH	75.95	17.17	150	4.00
TECC	63.86	13.83	153	5.41
PAC-MFCC	248.00	8.30	303	7.75
2 Sub-bandas ^a	422.40	15.48	907.8	22.40
4 Sub-bandas ^b	760.19	37.12	1473	37.80

^aPAC-MFCC com Classificador em Sub-Bandas utilizando 2 sub-bandas

^bPAC-MFCC com Classificador em Sub-Bandas utilizando 4 sub-bandas