



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CIN – Centro de Informática



PROPOSTA DE TRABALHO DE GRADUAÇÃO EM PROCESSAMENTO DE
VOZ

Reconhecimento de locutor utilizando HMM.

Aluno: Daniel Brito de Oliveira Carvalho (dboc@cin.ufpe.br)
Orientador: Tsang Ing Ren (tir@cin.ufpe.br)

Sumário

1.	Contexto	3
2.	Objetivo	4
3.	Cronograma	5
4.	Referências bibliográficas	6
5.	Assinaturas	8

➤ **Identificação**

Área de concentração

- Reconhecimento de Fala, Processamento de Sinais

Linha de pesquisa

- Inteligência Artificial

1. Contexto

A maioria de nós está consciente do fato de que vozes de indivíduos diferentes não soam iguais. É nesse contexto que a área de reconhecimento de locutor se insere, reconhecer pessoas através da voz. Logo reconhecimento de locutor é o processo de reconhecer pessoas usando informações específicas contidas no sinal de voz.

Sistemas de Reconhecimento de locutor podem operar em dois modos: identificação, identificar uma pessoa ou verificação, verificar se a pessoa é quem se diz ser. Esses sistemas podem ser classificados em sistemas dependentes do texto, no qual o reconhecimento está intimamente ligado ao texto e sistemas independente do texto.

Durante os ultimas décadas presenciou-se um grande avanço desses sistemas. Em 1974 Atal[1] alcançou taxas de erros próximas de 2%. Doddington em 1985[2] alcançou taxas de erros de 0,8%. Já em 1996 Colombi, et al[3] conseguiu taxas de erros próximas de 0,22%. Atualmente vemos essa taxa decaindo cada vez mais.

Diante dessa evolução uma das técnicas que se destacam é o HMM. o HMM, Hidden Markov Model, é um modelo estatístico que amplamente conhecido pela sua aplicação em reconhecimento de padrões temporais e vem sendo vastamente utilizado para o reconhecimento de locutor.

2. Objetivo

Este trabalho de graduação terá como objetivo analisar o reconhecimento de locutor utilizando HMM, Hidden Markov Models, enumerando seus pontos fracos e fortes. Será exposto todo o processo de criação de um sistema de reconhecimento de locutor utilizando HMM, desde a aquisição da base de dados, passando pela extração de características até chegar ao processo de treinamento e classificação.

Primeiramente, será feito o levantamento do material bibliográfico. Concomitantemente a esta etapa será adquirida ou criada uma base de dados. Em seguida, será realizada a implementação do sistema e por fim os resultados serão refinados e otimizados.

3. Cronograma

Atividades / Mês	Abril				Maio				Junho				Julho			
Levantamento Bibliográfico	■	■	■	■												
Aquisição e Escolha da Base de Dados			■	■	■											
Implementação				■	■	■	■	■	■	■						
Validação e otimização									■	■	■	■				
Escrita do relatório final							■	■	■	■	■	■	■			
Preparação da apresentação											■	■	■	■		

4. Referências bibliográficas

- [1] B. S. Atal, “Effectiveness of linear prediction characteristics of the speech wave for automatic speaker identification and verification,” *J. Acoust. Soc. Amer.*, vol. 55, no. 6, pp. 1304–1312, 1974
- [2] G. R. Doddington, “Speaker recognition—Identifying people by their voices,” *Proc. IEEE*, vol. 73, pp. 1651–1664, Nov. 1985.
- [3] J. Colombi, D. Ruck, S. Rogers, M. Oxley, and T. Anderson, “Cohort selection and word grammar effects for speaker recognition,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Atlanta, GA, 1996, pp. 85–88.
- [4] J. P. Campbell, “Speaker Recognition: A Tutorial.” *Proceedings of the IEEE*, Vol. 85, No.9, pp. 1437-1462, 1997

5. Avaliador

George Darmiton da Cunha Cavalcanti

6. Assinaturas

Daniel Brito de Oliveira Carvalho
Aluno

Tsang Ing Ren
Orientador