

Universidade Estadual Paulista
Instituto de Biociências, Letras e Ciências Exatas
Departamento de Ciência da Computação e
Estatística

Luis Fernando Teixeira Silva

Um sistema para reconhecimento de comandos falados
dependente do locutor

São José do Rio Preto - SP

2017

Luis Fernando Teixeira Silva

Um sistema para reconhecimento de comandos
falados dependente do locutor

Monografia apresentada ao Programa de
graduação em Ciência da Computação da
UNESP para obtenção do título de Bacharel.

Orientador: Prof. Dr. Rodrigo Capobi-
anco Guido

São José do Rio Preto - SP

2017

AUTORIZO A DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE

Ficha catalográfica elaborada pelo Serviço de Biblioteca do IBILCE/UNESP

Luis Fernando Teixeira Silva

titulo

titulo

titulo. / fulano de tal; orientador

Rodrigo Capobianco Guido. São José do Rio Preto, 2017.

xxx p.

Monografia (TCC

TCC

TCC, 2017.

1. Processamento de sinais. 2. Reconhecimento de locutor. 3. Acústica. 4. Energia. 5. Escala *Bark*. I. Capobianco Guido, Rodrigo, orient.
II. Título.

Dedico ao...

Agradecimentos

A Deus,

“No fim tudo dá certo, e se não deu certo é porque ainda não chegou ao fim.”

Fernando Sabino

Resumo

TAL, F. *titulo*. 2016. xxxp. TCC UNESP 2016.

Atualmente,

Palavras-chave: Processamento de sinais. Reconhecimento de locutor. Acústica. Escala *Bark*.

Abstract

TAL, F. *titulo*. 2016. xxxp. TCC UNESP 2016.

Nowadays, ...

Keywords: Signal processing. Speaker recognition. Acoustics. Bark scale.

Lista de Figuras

Lista de Tabelas

Lista de Abreviaturas

ABFF	Academia Brasileira de Fonoaudiologia Forense
CHMM	<i>Continuous Hidden Markov Models</i>
DTW	<i>Dynamic Time Warping</i>
DWT	<i>Discrete Wavelet Transform</i>
EER	<i>Equal Error Rate</i>
FIR	<i>Finite Impulse Response</i>
GMM-UBM	<i>Gaussian Mixture Models-Universal Background Models</i>
HMM	<i>Hidden Markov Model</i>
IEEE	<i>Institute of Electrical and Eletronic Engineers</i>
IIR	<i>Infinite Impulse Response</i>
LPCC	<i>Linear Predictive Cepstrum Coefficients</i>
MFCC	<i>Mel Frequency Cepstral Coefficients</i>
MLP	<i>Multiple Layer Perceptron</i>
NIST	<i>National Institute of Standards and Technology</i>
PCM	<i>Pulse Code Modulation</i>
RNA	Rede Neural Artificial
SVM	<i>Support Vector Machine</i>
TIMIT	<i>Texas Instruments and Massachusetts</i>
TTVNN	<i>Transição entre Trechos Vozeados e Não-Vozeados</i>
VQ	<i>Vector Quantization</i>
WAVE	<i>Waveform</i>
WPT	<i>Wavelet Packet Transform</i>
ZCR	<i>Zero Crossing Rate</i>

Sumário

1	Introdução	23
1.1	Considerações iniciais	23
1.2	Objetivos	23
1.3	Metodologia	23
1.4	Organização do trabalho	24
2	Revisão Bibliográfica	27
2.1	Fundamentação da Verificação de Locutores	27
2.2	Arquivos Acústicos no Formato <i>WAVE</i>	27
2.3	Energia	27
2.4	Vetores de Características	28
2.5	Níveis Críticos de Energia	28
3	Detalhamento do Trabalho Proposto	29
3.1	Considerações iniciais	29
4	Testes e Resultados	31
5	Conclusões e Trabalhos Futuros	33
	Referências	33
	Apêndice I - Gráficos das características extraídas	37

Capítulo 1

Introdução

1.1 Considerações iniciais

Petry (1) define o reconhecimento de locutores como...

1.2 Objetivos

Este trabalho tem como objetivo implementar um algoritmo computacional desenvolvido em C para reconhecer comandos falados dependente do locutor.

1.3 Metodologia

Para a elaboração deste projeto foi determinado os seguintes 11 comandos:

- Bom dia, Logan;

- Bom noite, Logan;
- Oi, Logan;
- Como está o tempo hoje?;
- vai chover?;
- Abrir calculadora;
- Ver notícias;
- Pesquisar;
- Alarme;
- Calendário;
- Sair;

sendo que posteriormente foi realizada a gravação de 10 áudios para cada um dos 11 comandos, totalizando 110 arquivos de áudio no formato *wave*.

1.4 Organização do trabalho

O texto vindouro do presente trabalho está organizado da seguinte forma:

- No Capítulo 2 apresenta-se uma série de trabalhos publicados envolvendo a área de reconhecimento de locutores, mostrando como são inúmeras as possibilidades de se realizar essa tarefa. Expõe-se, também, os principais conceitos e teorias que estão relacionados com o trabalho que foi desenvolvido.
- No Capítulo 3 apresenta-se, com detalhes, todo o desenvolvimento do trabalho proposto e de que forma os conceitos discutidos no capítulo anterior foram utilizados.

- No Capítulo 4 relatam-se todos os resultados obtidos no trabalho, a partir dos testes de reconhecimento que foram realizados.
- No Capítulo 5 apresentam-se as conclusões sobre o trabalho, bem como propostas para pesquisas futuras.

Capítulo 2

Revisão Bibliográfica

2.1 Fundamentação da Verificação de Locutores

Reconhecimento de locutores (??)...

2.2 Arquivos Acústicos no Formato *WAVE*

O formato...

2.3 Energia

A energia...

2.4 Vetores de Características

Vetor de característica...

2.5 Níveis Críticos de Energia

Níveis de energia são...

Capítulo 3

Detalhamento do Trabalho Proposto

3.1 Considerações iniciais

Como primeira etapa para

Capítulo 4

Testes e Resultados

bla bla bla...

Capítulo 5

Conclusões e Trabalhos Futuros

Neste trabalho, ...

Referências

- 1 PETRY, A. *Reconhecimento automático de locutor utilizando medidas de invariantes dinâmicas não-lineares*. 2002. 155 p. Tese (Doutorado em Ciência da Computação)-Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2002.
 - 2 CAMPBELL, J. P. et al. Forensic speaker recognition: a need for caution. *IEEE Signal Processing Magazine*, v. 26, n. 2, p. 95-103, 2009. doi:10.1109/msp.2008.931100.
 - 3 ACADEFORD. Disponível em: <<http://www.acadefor.com.br/>>. Acesso em: 12 ago. 2014.
- numérico. São Paulo: Pearson Prentice Hall, 2007.

Apêndice I - Gráficos das características extraídas