

RESUMO DA DISSERTAÇÃO APRESENTADA À UFPE COMO PARTE DOS REQUISITOS NECESSÁRIOS  
PARA OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA ELÉTRICA.

**ESTIMATIVAS DE COMPORTAMENTO VOCÁLICO DE LOCUTORES E UM NOVO  
SISTEMA DE SEPARAÇÃO SILÁBICA.**

**Elda Lizandra Fernandes da Silva.**

MAIO DE 2012

ORIENTADOR: PROF. DR. HÉLIO MAGALHÃES DE OLIVEIRA, DOCTEUR

ÁREA DE CONCENTRAÇÃO: TELECOMUNICAÇÕES/PROCESSAMENTO DE VOZ.

PALAVRAS-CHAVE: CARACTERIZAÇÃO DE VOZ, *PITCH*, SONS VOCÁLICOS. DIVISÃO SILÁBICA, LÍNGUA PORTUGUESA, PROCESSAMENTO DE FALA, CONVERSÃO FALA-PARA-TEXTO.

**NÚMERO DE PÁGINAS: 132**

Nesta dissertação um método simples para a estimação automática do comportamento espectral de trechos vocálicos de locutores é proposto. Uma implementação computacional em Matlab<sup>®</sup> é apresentada e sua validação é conduzida comparando os resultados com uma identificação realizada manualmente, empregando o Audacity 1.3<sup>®</sup>. Locutores (masculinos e femininos) foram considerados e os testes foram conduzidos para sete diferentes sons vocálicos da língua portuguesa (a, é, ê, i, ó, ô, u). A abordagem é potencialmente útil em modelos de trato vocal, na melhoria da qualidade de sintetizadores de voz ou em algoritmos de reconhecimento automático de locutor. Em uma segunda parte, um novo algoritmo para divisão silábica automática de arquivos de voz na língua portuguesa é proposto, com base na envoltória do sinal de voz. Uma implementação computacional em Matlab<sup>®</sup> é apresentada, a qual encontra-se disponibilizada na URL [http://www2.ee.ufpe.br/codec/divisao\\_silabica.html](http://www2.ee.ufpe.br/codec/divisao_silabica.html). Trechos longos contendo mais de uma sílaba e identificados com uma mesma envoltória são chamados de supersílabas e são separados posteriormente. Os resultados identificam as amostras correspondentes ao início e o fim de cada sílaba detectada. Foram realizados testes preliminares com meia centena de palavras, com uma taxa de identificação cerca de 70%, porém melhorias podem ser incorporadas para tratar fonemas nos quais o envelope não é o principal parâmetro na identificação. Este algoritmo também pode ser particularmente útil em sistemas com comandos de voz ou como ferramenta de apoio no ensino da língua portuguesa ou para pacientes em tratamento fonoaudiológico.

ABSTRACT OF DISSERTATION PRESENTED TO UFPE AS A PARTIAL FULFILLMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF MASTER IN ELECTRICAL ENGINEERING

## **Estimates of vowel speaker behavior and a new syllabic separation system**

**Elda Lizandra Fernandes da Silva.**

MAY 2012

ADVISOR: PROF. HÉLIO MAGALHÃES DE OLIVEIRA, DOCTEUR.

AREA OF CONCENTRATION: TELECOMUNICAÇÕES/VOICE PROCESSING

KEYWORDS: SPEECH CHARACTERIZATION, PITCH, VOWEL SOUNDS. SYLLABIC DIVISION,  
PORTUGUESE LANGUAGE, SPEECH PROCESSING, VOICE-TO-TEXT.

**NUMBER OF PAGES: 132**

In this thesis a new automatic method for estimating the spectral behavior of vowel excerpts of speakers is proposed. A computational implementation in Matlab<sup>TM</sup> is presented and the validation of the approach is conducted by comparing the results with an identification manually performed using Audacity 1.3<sup>TM</sup>. Speakers (both male and female) were considered and the tests were conducted for seven different vowel sounds of the Portuguese language (namely, a, é, ê, i, o, ô, u). The approach may be useful in models of the vocal tract, for improving the quality of speech synthesizers or to be used in algorithms for automatic speaker recognition. Also, a new algorithm for automatic syllabic splitting in the Portuguese language is proposed, which is based on the envelope of the speech signal of an audio file. A computational implementation in Matlab<sup>TM</sup> is presented and made available at the URL [http://www2.ee.ufpe.br/codec/divisao\\_silabica.html](http://www2.ee.ufpe.br/codec/divisao_silabica.html). Voice excerpts containing more than one syllable and identified by the same envelope are named as super-syllables and they are separated subsequently. The results indicate which samples corresponds to the beginning and end of each detected syllable. Preliminary tests were performed to fifty words at an identification rate circa 70% (further improvements may be incorporated to treat particular phonemes). This algorithm can also be particularly useful in voice command systems, as a tool in the teaching of Portuguese language or even for patients with speech pathology.