

Análise e Síntese de Fala

Este livro apresenta o desenvolvimento teórico e experimental de ferramentas de análise e síntese de sinais de fala. Nele são apresentadas e desenvolvidas várias técnicas de análise do sinal de fala com o objetivo de extrair um conjunto de parâmetros necessários à síntese da fala usando modelos de Formantes ou modelos LPC. O primeiro capítulo faz uma caracterização fonética da língua Portuguesa seguindo-se uma descrição dos cuidados a ter durante o processo de gravação de uma base de dados de fala. Depois é apresentado o desenvolvimento teórico das técnicas de análise Cepstral e análise por LPC's com as matrizes de autocorrelação e covariância. Estas técnicas são utilizadas para realizar o alisamento espectral e para extração automática das frequências formantes e frequência fundamental (F0). É ainda apresentada a implementação de um sintetizador de formantes empregue para testar os modelos desenvolvidos. Apresentam-se no capítulo 5 os detalhes de implementação de um sistema de conversão texto fala (Text-To-Speech) em Português.



João Paulo Teixeira

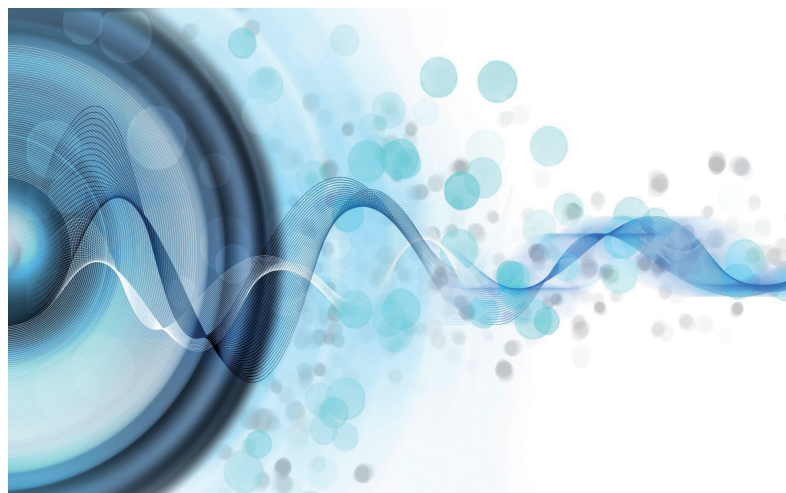
Doutorado pela FEUP em Engenharia Eletrotécnica e de Computadores, é professor no IPB na área de processamento de sinal. Pertence ao quadro editorial/comissão científica de algumas revistas e conferências científicas. É autor de dois livros, de alguns capítulos de livro e diversos artigos científicos sobre processamento de sinal e redes neuronais.



978-3-659-06206-3

editorial académica española

Análise e Síntese de Fala



João Paulo Teixeira

Análise e Síntese de Fala

Modelização Paramétrica de Sinais Para
Sistemas TTS

Teixeira

João Paulo Teixeira
Análise e Síntese de Fala

João Paulo Teixeira

Análise e Síntese de Fala

**Modelização Paramétrica de Sinais Para Sistemas
TTS**

Editorial Académica Española

Impressum / Impressão

Bibliografische Information der Deutschen Nationalbibliothek: Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Alle in diesem Buch genannten Marken und Produktnamen unterliegen warenzeichen-, marken- oder patentrechtlichem Schutz bzw. sind Warenzeichen oder eingetragene Warenzeichen der jeweiligen Inhaber. Die Wiedergabe von Marken, Produktnamen, Gebrauchsnamen, Handelsnamen, Warenbezeichnungen u.s.w. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutzgesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Informação biográfica publicada por Deutsche Nationalbibliothek: Nationalbibliothek numera essa publicação em Deutsche Nationalbibliografie; dados biográficos detalhados estão disponíveis na Internet: <http://dnb.d-nb.de>.

Os outros nomes de marcas e produtos citados neste livro estão sujeitos à marca registrada ou a proteção de patentes e são marcas comerciais registradas dos seus respectivos proprietários. O uso dos nomes de marcas, nome de produto, nomes comuns, nome comerciais, descrições de produtos, etc. inclusive sem uma marca particular nestas publicações, de forma alguma deve interpretar-se no sentido de que estes nomes possam ser considerados ilimitados em matérias de marcas e legislação de proteção de marcas e, portanto, ser utilizadas por qualquer pessoa.

Coverbild / Imagem da capa: www.ingimage.com

Verlag / Editora:

Editorial Académica Española

ist ein Imprint der / é uma marca de

AV Akademikerverlag GmbH & Co. KG

Heinrich-Böcking-Str. 6-8, 66121 Saarbrücken, Deutschland / Niemcy

Email / Correio eletrônico: info@eae-publishing.com

Herstellung: siehe letzte Seite /

Publicado: veja a última página

ISBN: 978-3-659-06206-3

Zugl. / Aprovado/a pela/pelo: Porto, Universidade do Porto, Dissertação, 1996

Copyright / Copirraite © 2013 AV Akademikerverlag GmbH & Co. KG

Alle Rechte vorbehalten. / Todos os direitos reservados. Saarbrücken 2013

ÍNDICE

LISTA DE ABREVIATURAS E SÍMBOLOS	VI
1. INTRODUÇÃO	1
1.1 Objetivos e Enquadramento Deste Trabalho	2
1.2 Aspetos de Organização	3
2. CARACTERIZAÇÃO FONÉTICA DA LÍNGUA PORTUGUESA	7
2.1 Introdução	8
2.2 Produção da Fala	8
2.2.1 Fonte Sonora	11
2.2.2 Frequência Fundamental	11
2.2.3 Formantes	12
2.3 Fonemas Do Português Europeu	13
2.3.1 Caracterização de Alguns Traços Fonéticos	13
2.3.2 Parâmetros Classificatórios das Consoantes	18
2.3.3 Lista de Fonemas do Português	21
3 RECOLHA DO SINAL DE FALA	23
3.1 Introdução	24
3.2 Elementos Importantes na Aquisição do Sinal de Fala	26
3.2.1 Câmara Insonorizada	26
3.2.2 Microfone	26
3.2.3 Pré-amplificador	26
3.2.4 Filtro Anti-Aliasing	26
3.2.5 Frequência de Amostragem	27
3.2.6 Número de Bits	28
3.2.7 Falante	28
3.2.8 Texto	28

3.3	Condições DE Aquisição dos Sinais de Fala	29
3.3.1	Câmara Insonorizada	30
3.3.2	Armazenamento do Sinal em Fita Magnética	30
3.3.3	Reprodução e Conversão Analógico/Digital com Recurso a Uma Placa de Som	30
3.3.4	Texto Escolhido	34
3.3.5	Falantes	34
4.	MODELIZAÇÃO PARAMÉTRICA DOS SINAIS DE FALA	35
4.1	Introdução	36
4.2	Modelos de Produção da Fala	37
4.2.1	Modelos de Tubos Sem Perdas	37
4.2.2	Modelo de Engenharia	38
4.2.3	Modelamento Sinusoidal da Fala	47
4.3	Parâmetros Para os Modelos dos Sinais de Fala	48
4.3.1	Parametrização por Formantes	48
4.4	Técnicas de Análise Baseadas no Modelo de Formantes	50
4.4.1	Análise por Síntese	50
4.4.2	Análise Síncrona com o Período Fundamental	51
4.4.3	Análise Cepstral	53
4.4.2	Análise por LPC's	55
4.5	Codificação Digital Com Qualidade de Telefonia da Forma de Onda da Voz	64
5.	CONVERSORES TEXTO-FALA	67
5.1	Introdução	68
5.2	Sistemas de Conversão Texto-Fala	69
5.2.1	Processamento Linguístico	69
5.2.2	Processamento Acústico	72
5.3	O Conversor Texto-Fala MULTIVOX	73
5.3.1	Blocos Constituintes	74
5.3.2	A Conversão Grafema Fonema Para o Português	78

5.3.3	Regras de Concatenação de Fonemas	83
5.3.4	Regras de Prosódia	84
6.	FERRAMENTAS DE SUPORTE À DETERMINAÇÃO AUTOMÁTICA DE PARÂMETROS	87
6.1	Introdução	88
6.2	Amplitude Média Deslizante	88
6.3	Energia Média Deslizante	91
6.4	Taxa de Passagem Por Zero	94
6.5	Segmentação/Classificação Quanto ao Modo de Excitação	96
6.6	Transformada de Fourier de Curta Duração (<i>Short-Time</i>)	105
6.7	Chirp <i>Z</i> Transform	106
6.8	Cepstro	108
6.9	espectrógrafo	112
6.10	Codificação por Predição Linear (LPC)	113
6.10.1	Matriz Autocorrelação	114
6.10.2	Matriz Covariância	115
6.11	Sintetizador	116
7.	DETERMINAÇÃO AUTOMÁTICA DOS PARÂMETROS DOS SINAIS DE FALA	119
7.1	Introdução	120
7.2	Determinação da Frequência Fundamental	121
7.2.1	Introdução	121
7.2.2	Método do Cepstro	121
7.2.3	Determinação da Frequência Fundamental no Erro Residual de Predição Linear	123
7.2.4	Processamento no Domínio Temporal	124
7.3	Determinação Automática dos Parâmetros Para a Fala Vozeada	131
7.3.1	Estrutura de Análise	131
7.3.2	Análise Síncrona Com o Período Fundamental	150
7.3.3	Alisamento Não Linear Aplicado à Sequência de Parâmetros Estimados	155

7.4	Análise de Fala Não Vozeada	157
8.	TESTE E AVALIAÇÃO DOS RESULTADOS	159
8.1	Introdução	160
8.2	Conversor Texto-Fala para o Português	160
8.3	Determinação da Frequência Fundamental com Processamento no Domínio Temporal	161
8.4	Extração Automática de Parâmetros dos Sinais de Fala	161
8.4.1	Método de Análise Cepstral	163
8.4.2	Método de Predição Linear - Matriz Autocorrelação	167
8.4.3	Método de Predição Linear - Matriz Covariância	171
8.4.4	Método de Análise Síncrona com o Período Fundamental	173
8.4.5	Comparação de Resultados dos Diferentes Métodos de Análise	176
9.	CONCLUSÃO	179
9.1	Conclusões	180
9.2	Outros Desenvolvimentos	183
	BIBLIOGRAFIA	185
	ANEXO A1	194
	ANEXO A2	196
	ANEXO A3	198
	ANEXO B1	206
	ANEXO B2	207
	ANEXO B3	208
	ANEXO B4	209
	ANEXO B5	211
	ANEXO B6	212

ANEXO B7	213
ANEXO B8	216
ANEXO B9	219
ANEXO B10	221
ANEXO B11	222
ANEXO B12	225
ANEXO B13	228
ANEXO B14	235
ANEXO B15	236
ANEXO B16	237
ANEXO B17	238
ANEXO B18	239
ANEXO B19	240
ANEXO B20	242
ANEXO B21	243

LISTA DE ABREVIATURAS E SÍMBOLOS

ABU - Acoustic Building Unit
ADM - Adaptative Delta Modulation
ADPCM - Adaptative Differential Pulse Code Modulation
APCM - Adaptive Pulse Code Modulation
CZT - Chirp Z Transform
DM - Delta Modulation
DPCM - Differential Pulse Code Modulation
 e - número de elementos abaixo de LSES
F0 - Frequência Fundamental
FFT - Fast Fourier Transform
Fi - formante de ordem i
Fimax - limite superior da região de frequências para o formante Fi
Fimin - limite inferior da região de frequências para o formante Fi
Fn - Frequência de Nyquist
Fs - Frequência de amostragem
G(z) - Função de transferência do impulso glotal
IDFT - Inverse Discrete Fourier Transform
LIZ - Limite Inferior da taxa de passagem por Zero
LSES - Limite Superior da Energia do Sinal
PCM - Pulse Code Modulation
PE – Português Europeu
R(z) - Função de transferência para o efeito de radiação nos lábios
RDP - Rádio Difusão Portuguesa
TPZ ou tpz - Taxa de Passagem por Zero
V(z) - Função de transferência do trato vocal
z - Número de elementos acima de LIZ

CAPÍTULO 1

INTRODUÇÃO

1. INTRODUÇÃO

1.1 OBJETIVOS E ENQUADRAMENTO DESTA TRABALHO

Neste livro apresenta-se o estudo e desenvolvimento de ferramentas de processamento de sinal que de uma forma automática realizam a extração de um conjunto de parâmetros de um sinal de fala com a menor intervenção possível do utilizador. Este conjunto de parâmetros deverá permitir que um sintetizador, baseado no mesmo modelo, alimentado por essa sequência de parâmetros produza novamente o sinal de fala original com o mínimo de distorção possível. Um sistema desta natureza é habitualmente designado por "VOCODER" e permite armazenar e transmitir sinais de fala ocupando, quer no armazenamento quer na transmissão, um número de bits muito menor que o sinal de fala original.

Um conversor texto-fala desenvolvido para o português Europeu faz uso de um conjunto de regras de concatenação de fonemas. Os fonemas são construídos pelo agrupamento de sons correspondentes a difones (duas metades de fonema) cortados nas suas zonas de sinal estacionário. O estabelecimento destas regras baseia-se num método denominado análise por síntese, em que se vai iterativamente chegando aos valores dos parâmetros por um processo de alteração de um valor destes, verificar o resultado e prosseguir as alterações no sentido de aproximar os sons sintetizados dos sons pretendidos. Este é um método de tentativa e erro um pouco moroso. Torna-se então clara a necessidade de um sistema que realize a análise de sinais de fala indicando assim quais as variações que devem seguir os parâmetros para produzir o som pretendido.

As ferramentas de análise aqui desenvolvidas têm aplicações em sistemas de análise de fala, em conversores texto-fala e sistemas de reconhecimento quer da fala quer do falante.

Estas matérias têm hoje um interesse crescente para aplicação na área da reabilitação, onde por exemplo um conversor texto-fala permite a um invisual trabalhar com recurso informáticos como o computador, que pronunciará o que se passa no ecrã, ou a um incapacitado, temporária ou permanentemente, da sua voz, recorrer a um programa adequado para fazer ouvir as suas mensagens. Por outro lado, os programas de análise permitem a um terapeuta da fala diagnosticar eventuais danos no aparelho fonador do paciente. Na área das comunicações estes sistemas têm o seu interesse na medida em que realizam uma compressão da informação dos sinais de fala permitindo reduzir a taxa de transmissão necessária para estes sinais (por ex. sistema GSM); na área da indústria existem já hoje