Pontifícia Universidade Católica
do Rio de Janeiro

Jan Krueger Siqueira

Reconhecimento de Voz Contínua com Atributos MFCC, SSCH e PNCC, Wavelet Denoising e Redes Neurais

DISSERTAÇÃO DE MESTRADO

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio.

Orientador: Prof. Abraham Alcaim

Rio de Janeiro
Setembro de 2011

Pontifícia Universidade Católica
do Rio de Janeiro

Jan Krueger Siqueira

Reconhecimento de Voz Contínua com Atributos MFCC, SSCH e PNCC, Wavelet Denoising e Redes Neurais

Dissertação apresentada como requisito parcial para obtenção do título de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica do Departamento de Engenharia Elétrica da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Abraham Alcaim
Orientador
Centro de Estudos em Telecomunicações - PUC-Rio

Fernando Gil Vianna Resende Junior
UFRJ

Prof. Marco Antonio Grivet Mattoso Maia
Centro de Estudos em Telecomunicações - PUC-Rio

Prof. José Eugênio Leal
Coordenador Setorial do
Centro Técnico Científico - PUC-Rio

Rio de Janeiro, 2 de setembro de 2011

Jan Krueger Siqueira

Graduou-se em Engenharia Elétrica nas áreas de Eletrônica e Telecomunicações pela PUC-Rio em 2008. Atua na empresa SmartSolutions Ltda, onde desenvolve sistemas para dispositivos móveis e para a web. Também ministra aulas de Matlab na PUC-Rio.

Ficha Catalográfica

Siqueira, Jan Krueger

Reconhecimento de voz contínua com atributos MFCC, SSCH e PNCC, wavelet denoising e redes neurais / Jan Krueger Siqueira ; orientador: Abraham Alcaim. – 2011.

85 f. ; 30 cm

Dissertação (mestrado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2011.

Inclui bibliografia

1. Engenharia elétrica – Teses. 2. Reconhecimento de voz contínua. 3. Ruído aditivo. 4. MFCC. 5. SSCH. 6. PNCC. 7. Wavelet denoising. 8. Rede neural. I. Alcaim, Abraham. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

A todos os professores que se dedicam de coração aos seus alunos.

Agradecimentos

Ao CNPq, por financiar a pesquisa.

Ao professores Abraham Alcaim e Marley Maria Bernardes Rebuzzi Vellasco, pelo ensino, inspiração e orientação neste projeto.

A Apple, Google, Mathworks e University of Cambridge, por fornecerem a tecnologia necessária.

À minha mãe, pelo amor e pelos ensinamentos desde que nasci.

A Paula Maurício Nunes, por ter sido uma grande companheira durante todo o mestrado. Mesmo não estando mais juntos, eu serei eternamente grato pelos dois anos de namoro e pela amizade.

Aos meus grandes amigos da PUC-Rio: Bruno Baère Pederassi Lomba de Araújo, Clarissa Costalonga e Gandour, Larissa Figueiredo Terra de Faria Reis, Vinícius Costa Villas Bôas Segura. Espero que a gente mantenha essa amizade especial que atravessou inúmeras aulas, almoços e cinemas.

Aos alunos, bolsistas e professores da minha turma de salsa do Jaime Aroxa. Os sábados continuam sendo especiais por causa de vocês.

A todos os familiares, pelo carinho de sempre.

A quem mais tenha facilitado, ainda que muito indiretamente, a conclusão deste trabalho.

Resumo

Siqueira, Jan Krueger; Alcaim, Abraham. **Reconhecimento de Voz Contínua com Atributos MFCC, SSCH e PNCC, Wavelet Denoising e Redes Neurais.** Rio de Janeiro, 2011. 85 p. Dissertação de Mestrado – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Um dos maiores desafios na área de reconhecimento de voz contínua é desenvolver sistemas robustos ao ruído aditivo. Para isso, este trabalho analisa e testa três técnicas. A primeira delas é a extração de atributos do sinal de voz usando os métodos MFCC, SSCH e PNCC. A segunda é a remoção de ruído do sinal de voz via wavelet denoising. A terceira e última é uma proposta original batizada de feature denoising, que busca melhorar os atributos extraídos usando um conjunto de redes neurais. Embora algumas dessas técnicas já sejam conhecidas na literatura, a combinação entre elas trouxe vários resultados interessantes e inéditos. Inclusive, nota-se que o melhor desempenho vem da união de PNCC com feature denoising.

Palavras-chave

Reconhecimento de voz contínua; ruído aditivo; MFCC; SSCH; PNCC; wavelet denoising; rede neural.

Abstract

Siqueira, Jan Krueger; Alcaim, Abraham. **Continuous Speech Recognition with MFCC, SSCH and PNCC Features, Wavelet Denoising and Neural Networks.** Rio de Janeiro, 2011. 85 p. M. Sc. Dissertation – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

One of the biggest challenges on the continuous speech recognition field is to develop systems that are robust to additive noise. To do so, this work analyses and tests three techniques. The first one extracts features from the voice signal using the MFCC, SSCH and PNCC methods. The second one removes noise from the voice signal through wavelet denoising. The third one is an original one, called feature denoising, that seeks to improve the extracted features using a set of neural networks. Although some of these techniques are already known in the literature, the combination of them brings many interesting and new results. In fact, it is noticed that the best performance comes from the union of PNCC and feature denoising.

Keywords

Continuous speech recognition; additive noise; MFCC; SSCH; PNCC; wavelet denoising; neural network.

Sumário

Lista de Figuras

Lista de Tabelas

Your time is limited, so don't waste it living someone else's life. Don't be trapped by dogma — which is living with the results of other people's thinking. Don't let the noise of others' opinions drown out your own inner voice. And most important, have the courage to follow your heart and intuition. They somehow already know what you truly want to become. Everything else is secondary.
Steve Jobs

It's in Apple's DNA that technology alone is not enough. That it's technology married with liberal arts, married with the humanities, that yields us the result that makes our hearts sing.
Steve Jobs

Dance bem, dance mal, dance sem parar!
Nelson Motta / Rubens Queiroz