

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Verificação de locutor utilizando parâmetros extraídos do sinal glotal em conjunto com a técnica MFCC

Carla F. Schueler¹

Departamento de Engenharia de Telecomunicações, UFF, Niterói, RJ

Filipe Silveira²

Departamento de Engenharia de Telecomunicações, UFF, Niterói, RJ

Edson Cataldo³

Departamento de Matemática Aplicada, Programa de Pós-graduação em Engenharia Elétrica e de Telecomunicações, UFF, Niterói, RJ

1 Introdução

A voz é um fenômeno físico resultante da propagação do fluxo de ar, proveniente dos pulmões, através da glote e do trato vocal, com posterior irradiação pela boca. A unicidade dos órgãos envolvidos no processo de geração da voz é o que torna a fala um sinal biométrico, podendo ser usada para a identificação de um locutor. Sistemas de verificação e reconhecimento de locutor podem ser então construídos para os mais diversos fins, como por exemplo, para sistemas que utilizam a voz como chave de acesso a ambientes restritos. Esse artigo tem por objetivo tratar de um algoritmo de verificação de locutor, aliando, para isso, coeficientes MFC (mel frequency cepstrum) extraídos de sinais de voz, com características extraídas do sinal glotal, que é um sinal obtido imediatamente após a passagem do ar pelas cordas vocais. O sinal glotal é obtido por um processo de filtragem inversa. Acredita-se que o sinal glotal carregue informações importantes das características do locutor, pois ainda não sofreu a influência do trato vocal, possibilitando, assim, a eficiência do sistema proposto.

2 Metodologia

O algoritmo de verificação de locutor desenvolvido utiliza uma base de áudios com amostras de um mesmo locutor, das quais são extraídos os parâmetros mencionados que, por sua vez, são usados para treinar o modelo característico do locutor de interesse, a partir da aplicação da teoria de modelos ocultos de Markov. Áudios de teste são submetidos ao modelo, gerando como resposta um vetor de verossimilhanças correspondentes às amostras

¹carla_schueler@id.uff.br

²filipesilveira@id.uff.br

³ecataldo@im.uff.br

testadas. Para testar a ferramenta de verificação, foi desenvolvida uma técnica capaz de julgar o quão próximo do modelo gerado um grupo de áudios deveria estar, para ser interpretado como pertencente ou não ao locutor de interesse. O mecanismo de decisão adota como base a distância entre as funções densidade de probabilidade do logaritmo das verossimilhanças, ilustradas na Fig. 1. Áudios testes do locutor de interesse foram submetidos ao modelo e, a partir do logaritmo das verossimilhanças resultantes, criou-se a chamada curva padrão. Calculou-se, então, a distância entre as curvas resultantes dos testes de áudios de locutores desconhecidos pelo modelo e a curva padrão.

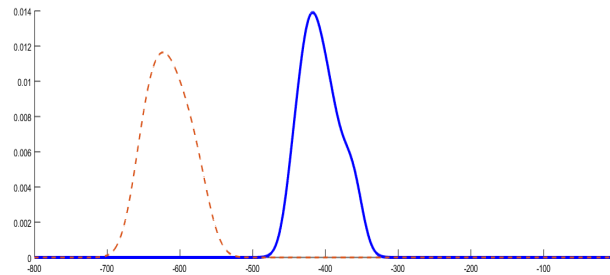


Figura 1: Funções densidade de probabilidade das verossimilhanças: locutor de interesse (linha cheia-curva padrão) e locutor desconhecido (linha tracejada).

3 Resultados e conclusões

Os resultados apontam que as curvas dos locutores desconhecidos, utilizando a técnica MFCC em conjunto com os parâmetros do sinal glotal, afastam-se da curva padrão, em relação aos resultados obtidos apenas quando a técnica MFCC é utilizada. Esse resultado permite que o algoritmo de verificação de locutor desenvolvido aceite mais vezes os áudios do locutor de interesse e recuse completamente os áudios dos locutores desconhecido, caracterizando a robustez do sistema. Tal fato é inovador, pois apesar da ideia que os parâmetros extraídos do sinal glotal carregam mais informações sobre o locutor, pouco existe na literatura sobre a aplicação dessa técnica a fim de melhorar processos de verificação de locutor como o desenvolvido.

Referências

- [1] L. F. Mendoza, E. Cataldo, M. B. R. Vellasco, M. A. Silva and A. J. Apolinário. Classification of Vocal Aging Using Parameters Extracted From the Glottal Signal, *Journal of Voice*, 28:532–537, 2014.
- [2] K. Murphy. *Hidden Markov Model (HMM) Toolbox for Matlab*, <http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html>, 1988.
- [3] I. Titze. *Principles of Voice Production*. PRT, Prentice-Hall, 1993.