

Sistema de Extração de Características Cíclicas de Vozes Patológicas utilizando *App Designer*

Marisa F. Ciríaco¹, Antônio J. B. Freitas¹, Maurício F. Fernandes¹, Aluisio I. R. Fontes¹

¹Instituto Federal do Rio Grande do Norte (IFRN)
– 59.900-000 – Pau dos Ferros – RN – Brazil

{marisa.franca, jhoanes.freitas, mauricio.ferreira}@escolar.ifrn.edu.br,
aluisio.rego@ifrn.edu.br

Abstract. *This work proposes to demonstrate how the MATLAB App Designer can be useful in processing digital signal research and its application to extract features in pathological voices. Cyclic correlogram spectral density (CCSD) was used to extract features from the Massachusetts Eye and Ear Infirmary (MEEI) database. The results demonstrated the potential of the CCSD as pathological voices descriptor and App Designer as an innovative tool for graphical interfaces in MATLAB.*

Resumo. *Este trabalho propõe demonstrar como o App Designer do MATLAB pode ser útil em pesquisas de processamento digital de sinais e sua aplicação para extração de características em vozes patológicas. A densidade espectral de correlograma cíclica (CCSD) foi utilizada para extrair características no banco de dados Massachusetts Eye and Ear Infirmary (MEEI). Os resultados demonstraram o potencial da CCSD como descritor de vozes patológicas e o App Designer como uma ferramenta inovadora para interfaces gráficas no MATLAB.*

Introdução

Com o decorrer do pós-modernismo, o risco de danos às pregas vocais tem crescido por variados fatores relacionados ao estilo de vida, tais como a poluição, agentes patológicos, tabagismo e sobrecarga de trabalho. Pelo menos 25% da população mundial utiliza sua voz como meio de vida, por exemplo, professores, cantores, atores, advogados e supervisores [Al-nasheri et al. 2017]. Analisando a prevalência de distúrbios vocais nos EUA, estudos mostraram que, para professores, cerca de 57,7% dos profissionais apresenta algum transtorno vocal durante a vida, enquanto profissionais de outras áreas tiveram uma incidência de 28,8% [Roy et al. 2004]. Por a voz humana ser um importante meio de comunicação, qualquer perturbação vocal pode ter consequências negativas na vida profissional e social do indivíduo [Almeida 2010].

Nos últimos anos, os avanços tecnológicos vêm contribuindo significativamente na área médica, permitindo que cirurgias, exames e diagnósticos sejam mais confiáveis e cada vez menos invasivos. Em particular, técnicas de análise de sinais digitais surgem como uma alternativa eficaz, não invasiva e de baixo custo para detecção de patologias na laringe [Cordeiro et al. 2016]. Isso é possível porque as doenças da voz influenciam as vibrações das cordas vocais diferentemente, dependendo da anomalia e do local que está sendo afetado, fazendo com que cada patologia reproduza tons específicos particulares a ela, que podem ser identificados e classificados através dessas técnicas [Costa 2008].

O desenvolvimento de novas tecnologias de processamento e análise digital de sinais, se mostra, então, essencial tanto no processo de diagnóstico como no acompanhamento terapêutico do paciente. A análise acústica desses sinais é uma alternativa complementar aos métodos de inspeção direta das cordas vocais, o que pode diminuir a frequência de exames invasivos [Fontes et al. 2014]. As técnicas existentes de processamento, geralmente, envolvem várias fases de pré-processamento de uma série de bancos de dados de voz, onde os pacientes reproduzem um fonema qualquer que possa ser falado confortavelmente, como por exemplo, a vogal /a/. Depois desse processo, em nosso estudo, o sinal passará por um extrator de características de baixa complexidade computacional e, posteriormente, um classificador. Métodos tradicionais utilizam análise estatística de segunda ordem para extrair características da voz. Esse trabalho utilizou a CCSD proposta por [Fontes et al. 2017] para extrair características cicloestacionárias utilizando infinitos momentos estatísticos, aplicando diferentes valores para o Kernel, que é uma variável livre da CCSD definida pelo usuário. Para esse estudo, através de resultados empíricos, utilizamos Kernel no valor de 0,1. Em trabalhos futuros, será estudado e definido o melhor valor de Kernel para a extração de características mais eficiente possível.

Portanto, este trabalho tem como objetivo mostrar como a nova ferramenta *App Designer* do MATLAB (versão R2016b) pode ser útil no desenvolvimento de *softwares* para o processamento digital de sinais, ajudando graficamente a compreensão de alguns algoritmos. Também será analisado o desempenho da densidade espectral de correntropia cíclica na extração de características do sinal da voz. Resultados preliminares demonstraram que os infinitos momentos estatísticos de ordem par da CCSD são capazes de extrair características distintas nas vozes patológicas. Esses resultados foram gerados a partir do *App Designer* para facilitar a compreensão de novos pesquisadores.

Este trabalho foi organizado da seguinte forma: a sessão 2 aborda um pouco sobre o *App Designer* e sua utilidade para pesquisas de processamento digital de sinais, na sessão 3 são mostrados alguns resultados obtidos com extração de características de vozes, e a conclusão do trabalho é dada na sessão 4.

***App Designer* MATLAB**

O *App Designer* é um ambiente de criação gráfica desenvolvido recentemente, a partir da versão R2016a, para aprimorar seu antecessor, o GUIDE. A nova versão do ambiente gráfico deixa o código mais legível, e consideravelmente mais simples de ser programado. Por exemplo, enquanto o código gerado pela GUIDE é completamente editável, no *App Designer*, só é permitido ao pesquisador que edite *callbacks* (chamadas de eventos ao interagir com determinado componente), algumas funções e declarações criadas pelo usuário.

A parte do código não editável é indicada por um fundo cinza. Isso evita que o código gerenciado pelo MATLAB na interface criada seja sobrescrito acidentalmente, levando a falhas no programa. Além disso, o *App Designer* apenas gera *callbacks* se o usuário explicitamente requisitar, impedindo que *callbacks* vazios acabem sobrecarregando e desorganizando o código. Ademais, para que o código se torne mais legível, a estrutura é dividida em várias sessões nessa ordem: Propriedades da aplicação, funções definidas pelo usuário, e o código usado para criação e inicialização de componentes.

Ambiente gráfico

O *App Designer* simplifica o processo de criação de componentes visuais para uma interface de usuário. Dessa forma, ele inclui um conjunto completo dos principais componentes comerciais, que ficam disponíveis em uma área de *design*, contendo botões, interruptores e demais itens para montagem de painéis de controle que aperfeiçoam a interação do usuário com os algoritmos e otimizam o tempo levado para execução de testes essenciais na pesquisa. Adicionalmente, o *App Designer* integra as duas principais formas de criação de aplicativos: por codificação direta no editor do código da interface criada (aba "*code view*") e por codificação indireta, arrastando componentes dentro do ambiente de criação (aba "*design view*"). O editor incorporado proporciona que o desenvolvedor possa adicionar novas propriedades, *callbacks* e funções compartilhadas de forma simples e intuitiva.

O ambiente gráfico do *App Designer*, Figura 1, é composto pela tela de componentes (1), os quais podem ser arrastados até a tela de design (*design view*), enquanto na tela de propriedades especializadas (3) pode-se definir configurações específicas para os componentes. Além disso, é possível usar as dicas de alinhamento para obter um *layout* preciso dos componentes (2) e definir as propriedades dos eixos no ambiente de projeto. Caso deseje, o pesquisador poderá escrever o próprio código. Para isso, basta que seja acessado a aba "*Code View*", Figura 2.

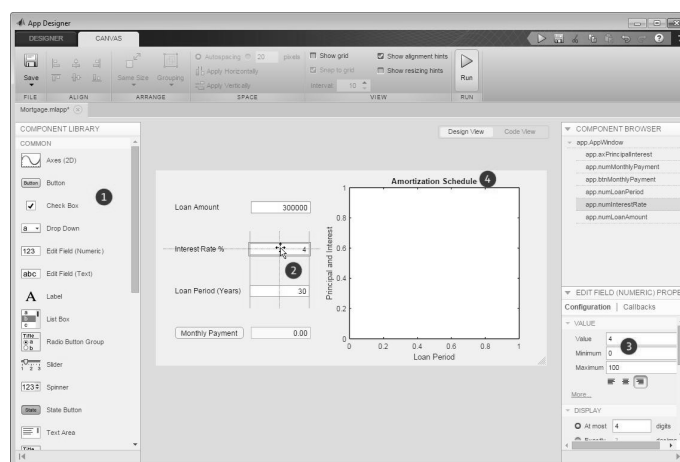


Figura 1. Interface Gráfica *App Design - Design View*

App Designer e Processamento digital de Sinal

Por ser uma ferramenta completa de engenharia, o MATLAB fornece uma vasta gama de opções de manipulação de sinais. Possui implementação para utilizar inteligência artificial, possuindo suporte para redes neurais, tornando possível o reconhecimento de padrões. Com o *App Designer* podemos criar componentes que auxiliam a manipulação dos dados desses sinais de forma prática. Por exemplo, na interface criada em nosso estudo, foi utilizado um botão que altera diretamente o valor de Kernel na função de retirada de características da voz; assim é possível testar como as frequências se comportam nessas variações sem que seja preciso alterar o código manualmente. Ademais, foi possível a criação de *checkboxes* (caixas de seleção) para se ter a opção de escolher qual tipo de

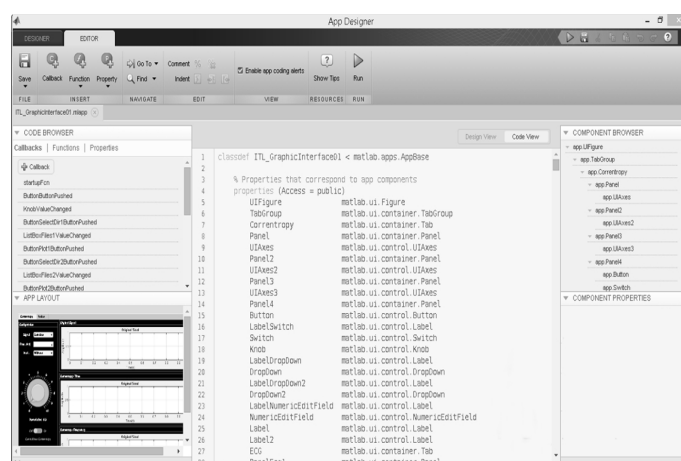


Figura 2. Interface Gráfica App Design - Code View

gráfico será exibido. Em cada *checkbox* é factível marcar qual opção de filtro queremos testar no momento.

Sem essa funcionalidade teríamos que, na necessidade de testar diversas variações possíveis, alterar o código diretamente para que o compilador pudesse ignorar ou não essas instruções. As funções próprias do *App Designer*, tanto para leitura e processamento digital de sinais, como para manipulação de arquivos, podem ser consultadas detalhadamente numa documentação extensa e didática, ou através de um simples comando de ajuda na janela de comando do MATLAB. Por exemplo, se precisa saber informações detalhadas sobre a função “audioread”, basta digitar “help audioread” na janela de comando (sem aspas).

O *App Designer* também possui diversas opções de demonstração de gráficos que proporcionam, por exemplo, comparação entre várias ondas de sinal em uma única janela, visão detalhada com possibilidade de *zoom* (lente de aumento) e a possibilidade de mostrar várias janelas de gráficos diferentes simultaneamente. Na sessão de resultados (capítulo 3) as Figuras 4, 6 e 8 mostram gráficos dos sinais de voz em função do tempo, e as Figuras 5, 7 e 9 mostram a voz passando pelo extrator de características (CCSD). As vozes mostradas são uma normal, uma com edema e uma com nódulo. Segue abaixo uma imagem da interface criada, onde temos um botão de seleção de diretório, para selecionarmos onde se encontram os sinais, que serão mostrados em uma lista, podendo ser selecionados individualmente para processamento digital utilizando as opções de demonstração de gráfico da frequência da onda em função do tempo (plot no tempo), ou passando pelo extrator de características (CCSD), Figura 3.

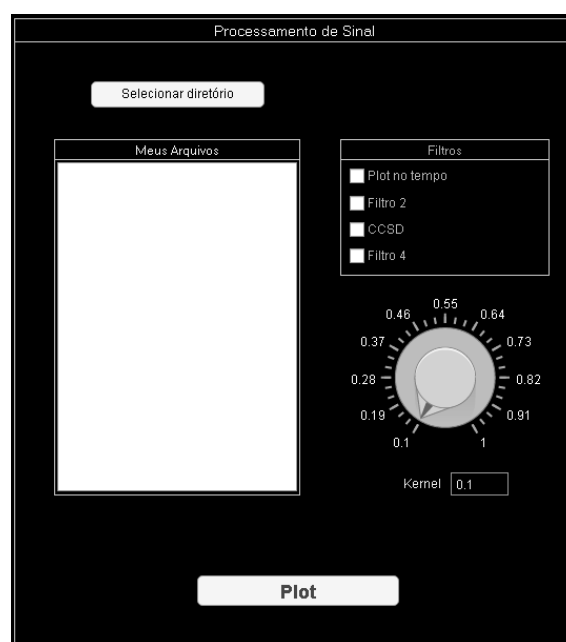


Figura 3. Interface Gráfica *App Design*

Resultados

Neste estudo foi usado o banco de dados Massachusetts Eye and Ear Infirmary (MEEI) Voice and Speech. Esse banco de dados possui cerca de 1400 amostras e é vastamente utilizado em estudos de identificação e classificação de patologias na voz. Nesse banco de dados os pacientes com vozes patológicas possuem desordens de caráter orgânico, traumático e psicogênico [Amami and Smiti 2016]. Para nosso estudo, utilizamos três vozes: uma voz normal, uma com edema, e uma com nódulo. Todas as pessoas passaram por procedimentos médicos de avaliação das condições da laringe, tais como estroboscopia e exames físicos da região do pescoço e boca. Todas as amostras do banco de dados foram gravadas sustentando-se a vogal /a/ e a primeira parte da passagem do arco-íris. A frequência das amostras para vozes normais foi de 50 kHz, enquanto as vozes patológicas 50 kHz ou 25 kHz [Al-nasheri et al. 2017].

Segue abaixo uma demonstração dos resultados obtidos ao gerar os gráficos dos sinais de voz, e extração de características. Os gráficos das Figuras 4, 6 e 8 mostram a frequência do sinal da voz em função do tempo em segundos, e os gráficos nas Figuras 5, 7 e 9 mostram as assinaturas de cada voz, ao passarem pelo extrator de características (CCSD). Ao olhar para os gráficos de extração de características (Figuras 5, 7 e 9), percebe-se que cada tipo de voz possui padrões distintos peculiares a ela, permitindo que o procedimento de classificação seja possível com eficiência, detectando esses padrões para que a voz possa ser categorizada em determinado grupo, patológico ou não. Nos gráficos, a variável "K" representa o Kernel.

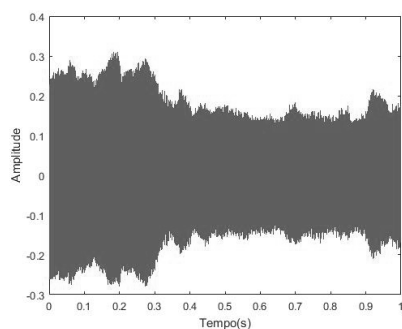


Figura 4. Voz normal JKR25 plotada em função do tempo em segundos

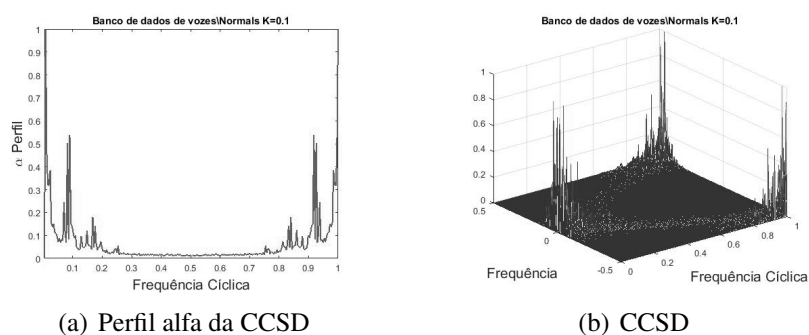


Figura 5. Extração de características da voz normal JKR25

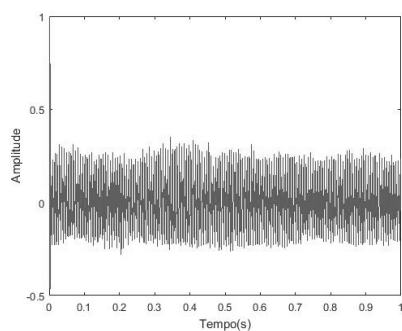


Figura 6. Voz com edema ANA15AN plotada em função do tempo em segundos

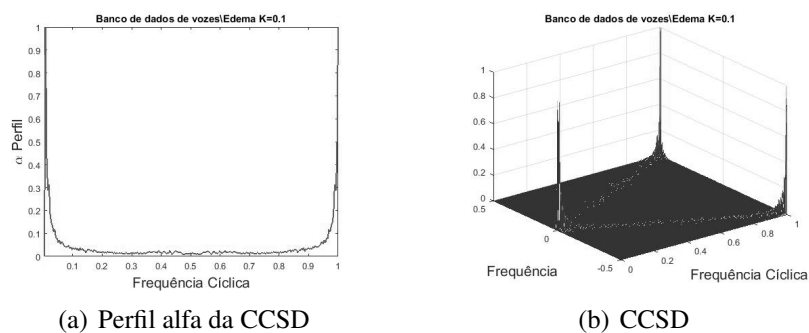


Figura 7. Extração de características da voz com edema ANA15AN

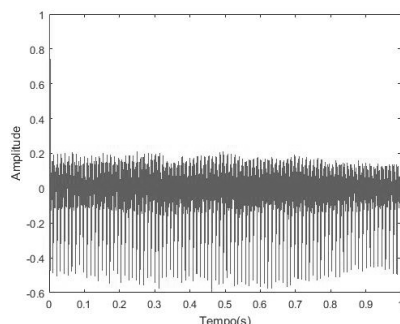


Figura 8. Voz com nódulo AMC14AN plotada em função do tempo em segundos

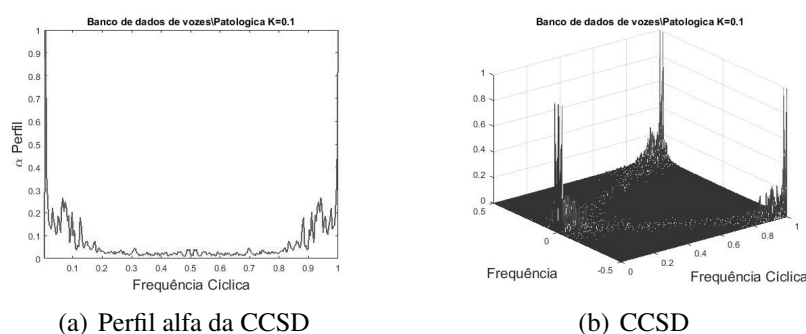


Figura 9. Extração de características da voz com nódulo AMC14AN

Conclusão

O presente artigo demonstrou que o uso do *App Designer* do MATLAB é eficaz na criação de uma interface gráfica, que pode diminuir significativamente o tempo consumido com testes, permitindo ao pesquisador dedicar-se mais aos resultados da pesquisa, sem que seja demandado muito tempo para a programação dos componentes e funções que serão aplicados ao processamento dos sinais. Por não ser uma aplicação complexa, e sua interação com o usuário ser intuitiva e de fácil aprendizado, pode contribuir de forma notável para pesquisas e desenvolvimento de aplicativos em áreas, outrossim, onde os pesquisadores não possuem necessariamente um conhecimento aprofundado de programação. Além disso, os resultados experimentais demonstraram que a CCSD é uma ferramenta eficaz para extração de características. Diante do trabalho proposto, podem ser apontadas as seguintes propostas de trabalhos futuros: Implementação da etapa de classificação, desenvolvimento de novos filtros e algoritmos e extração de características de novos sinais biológicos.

Referências

- Al-nasheri, A., Muhammad, G., Alsulaiman, M., and Ali, Z. (2017). Investigation of voice pathology detection and classification on different frequency regions using correlation functions. *Journal of Voice*, 31(1):3–15.
- Almeida, N. C. d. (2010). Sistema inteligente para diagnóstico de patologias na laringe utilizando máquinas de vetor de suporte.

- Amami, R. and Smiti, A. (2016). An incremental method combining density clustering and support vector machines for voice pathology detection. *Computers & Electrical Engineering*.
- Cordeiro, H., Fonseca, J., Guimarães, I., and Meneses, C. (2016). Hierarchical classification and system combination for automatically identifying physiological and neuromuscular laryngeal pathologies. *Journal of Voice*.
- Costa, S. L. d. N. C. (2008). *Análise Acústica, Baseada no Modelo Linear de Produção da Fala, para discriminação de vozes patológicas*. PhD thesis, Universidade Federal de Campina Grande.
- Fontes, A. I., Rego, J. B., Martins, A. d. M., Silveira, L. F., and Principe, J. (2017). Cyclostationary correntropy: Definition and applications. *Expert Systems with Applications*, 69:110–117.
- Fontes, A. I., Souza, P. T., Neto, A. D., Martins, A. d. M., and Silveira, L. F. (2014). Classification system of pathological voices using correntropy. *Mathematical Problems in Engineering*, 2014.
- Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., and Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research*, 47(2):281–293.