

Avaliação de Algoritmo para Detecção da Frequência Fundamental em Sinais de Vibração da Pele do Pescoço

Lucas M. Souza (IFPB, Campus João Pessoa), Igor F. Silva (IFPB, Campus Itabaiana), Thyago L. V. Silva (IFPB, Campus João Pessoa) e Suzete E. N. Correia (IFPB, Campus João Pessoa)

E-mails: souza.mendes@academico.ifpb.edu.br, igor.silva@ifpb.edu.br, thyago.lima@ifpb.edu.br, suzete@ifpb.edu.br.

Área de conhecimento (Tabela CNPq): 3.13.01.01-0 Processamento de Sinais Biológicos.

Palavras-chave: processamento de sinais, qualidade vocal, tecnologia vestível, sistema embarcado.

1. Introdução

A frequência fundamental do sinal de voz, ou *pitch*, reflete o número de ciclos vibratórios das pregas vocais por segundo. A taxa de vibração é determinada principalmente pela massa e tensão, das pregas vocais, embora a pressão e velocidade do ar também contribuam, em menor escala. Alterações na frequência fundamental prejudicam a qualidade vocal e podem ser devidas a causas orgânicas (infecções, cistos e pólipos benignos); neurológicas; funcionais (uso inapropriado da voz) e mecânicas (alterações fisiológicas na laringe) (BEHLAU, 2001).

A análise acústica, por meio de técnicas de processamento digital de sinais, possibilita a avaliação objetiva e quantitativa da qualidade vocal, identificação de patologias laringeas e monitoramento da evolução do tratamento. Uma das medidas acústicas que podem ser detectadas, por meio dessa análise, é a frequência fundamental (PRUKKANON et al., 2007). Embora seja um método não invasivo e uma ferramenta que contribui para o estudo da fisiologia vocal, está sujeita a ruídos e sons ambientais, durante a captura do sinal de voz por um microfone, o que pode afetar o seu desempenho.

Outras tecnologias têm sido empregadas para mensurar sinais derivados da voz, a exemplo das que usam sensores como acelerômetros e pastilhas piezoelétricas, posicionados no início do osso esterno no pescoço, para a captura de sinais de Vibração da Pele do Pescoço (VPP) (CHWALEK et al., 2018; SILVA et al., 2019). Tal análise é robusta a ruídos presentes no ambiente e fornecem sinais propensos a melhor detecção de fonemas sonoros, por estarem sendo extraídos próximo à fonte glótica.

Como os sinais de VPP possuem informações sobre a produção da voz é possível extrair parâmetros, que podem quantificar a eficiência e a qualidade do sistema vocal do locutor. Chwalek et al. (2018) propuseram um dispositivo vestível compacto, em forma de colar e com conexão sem fio para monitoramento dos parâmetros vocais. Utilizaram como sensor, um acelerômetro de efeito piezoelétrico, que deve ficar em contato com o pescoço. Com este estudo, foi possível detectar a frequência fundamental, o nível de pressão sonora e o tempo de fonação.

Silva et al. (2019) propuseram um colar de baixo custo para capturar os sinais de vibração da pele do pescoço empregando um sensor piezoelétrico de 35 mm de diâmetro. Um microfone foi empregado para realizar uma comparação com a análise acústica, de modo a avaliar o funcionamento do dispositivo desenvolvido. A frequência fundamental, obtida com o uso do dispositivo apresentou valores análogos aos valores obtidos por meio da análise acústica. No entanto, as medidas foram extraídas por meio do PRAAT®, um software gratuito e de código aberto, utilizado para análise e síntese da fala.

Para desenvolvimento de um sistema que permita o monitoramento remoto e em tempo real das vibrações da pele do pescoço, torna-se necessário que a detecção da frequência fundamental seja realizada em um dispositivo embarcado. Para realizar essa tarefa há diversos algoritmos capazes de detectar a frequência fundamental de sinais de áudio. Podem ser encontrados algoritmos que trabalham no domínio do tempo, como o *Average Magnitude Difference Function* (AMDF) e baseados na autocorrelação, como também métodos no domínio da frequência, que empregam a Transformada Rápida de Fourier (FFT – *Fast Fourier Transform*), análise cepstral e máxima verossimilhança (PRUKKANON et al., 2007, RABINER et al. 1976).

O presente trabalho apresenta o resultado da implementação de uma técnica baseada no domínio da frequência, que foi embarcada em uma placa ESP32, para a detecção da frequência fundamental de sinais de VPP. O erro percentual médio e o custo computacional foram considerados para avaliação de desempenho.

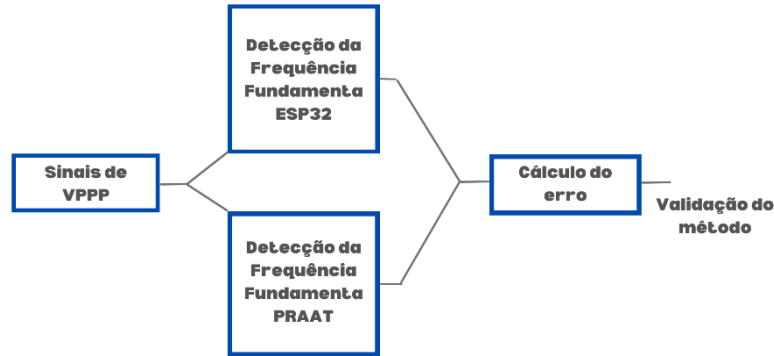
2. Materiais e métodos

A Figura 1 apresenta o diagrama em blocos da metodologia empregada. Dos sinais de vibração da pele do pescoço é extraída a frequência fundamental por meio de um método baseado na FFT, implementado na linguagem C e embarcado no microcontrolador ESP32. Para fins de comparação e cálculo do erro percentual foi obtida também a frequência fundamental empregando-se o software PRAAT®.

Os sinais de VPP fazem parte de uma base de dados aprovada pelo Comitê de Ética em Pesquisa (CEP), do IFPB, de nº 3.577.570 (SILVA et al., 2019). Essa base possui 128 sinais, de 32 voluntárias. Para este trabalho são consideradas as

emissões da vogal sustentada /ε/, na intensidade normal. A frequência de amostragem foi de 8000 Hz, sendo 12 bits por amostra.

Figura 1 – Diagrama em blocos da metodologia.



Fonte: Autoria própria.

Para a detecção da frequência fundamental dos sinais de VPP, inicialmente é obtida a sua representação no domínio da frequência, por meio da FFT, como ilustrado na Figura 2. O propósito do algoritmo é identificar o pico de menor valor em frequência que corresponda à frequência fundamental. Não necessariamente esse pico possui o máximo valor global do espectro. A não utilização de um método de busca eficiente, levará a respostas erradas. Na busca por um método eficiente, foram propostas as seguintes etapas a serem executadas:

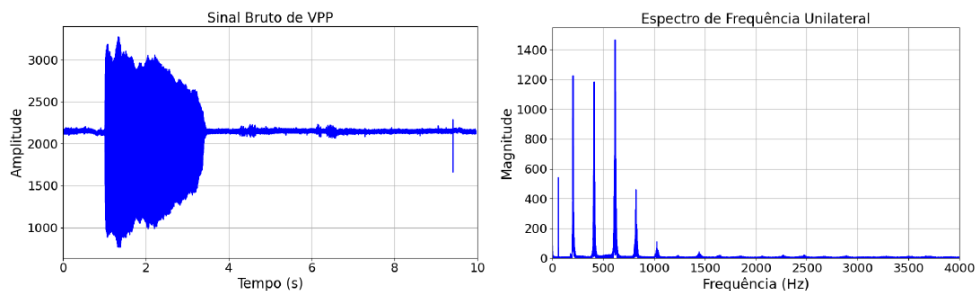
1. *Definição da faixa de interesse no espectro:* o processamento ocorre apenas em uma faixa específica de frequência dentro da FFT. A faixa utilizada durante os experimentos foi de 100 Hz a 900 Hz. A magnitude e a frequência de cada ponto dentro dessa faixa foi armazenado em um estrutura de dados como um par ordenado (x,y).

2. *Ordenação por magnitude:* a estrutura de dados foi ordenada de modo decrescente utilizando a magnitude como critério de ordenação, de modo que os picos mais altos da FFT apareçam primeiro. O algoritmo de ordenação utilizado foi o heapsort, que possui um tempo computacional relativamente baixo.

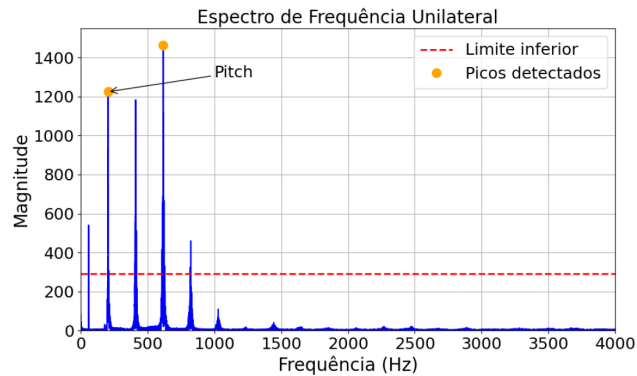
3. *Identificação do pico dominante com espaçamento mínimo:* o primeiro elemento da estrutura de dados do espectro ordenado é o pico global da FFT. Essa etapa começa considerando esse primeiro elemento como o pico correspondente à frequência fundamental. Esse elemento pode ser de fato o fundamental do sinal ou pode ser, provavelmente, um outro harmônico que possui uma magnitude elevada (como pode ser visto na Figura 3). Então, os elementos da estrutura são percorridos em ordem decrescente de magnitude e caso seja encontrado um elemento onde sua ocorrência no espectro original é anterior ao pico guardado atualmente e, a distância entre o pico atual e esse elemento é maior ou igual a um valor de 1 em Hz, esse elemento é guardado como o novo pico atual. Os elementos da estrutura são percorridos até alcançar um limite inferior de magnitude, como pode ser visto na Figura 3. Nos experimentos esse limite inferior foi definido como 20% da magnitude máxima encontrada na estrutura de dados (ou seja, do primeiro elemento da estrutura). Em suma, essa etapa busca encontrar o primeiro pico da FFT que ocorra dentro de uma faixa específica e que possua uma magnitude superior a um limite inferior predefinido.

Após a detecção da frequência fundamental é calculado o erro percentual entre o valor obtido pelo algoritmo e o valor fornecido pelo software PRAAT®, para validação do método, através da Equação 1.

Figura 2 – Sinal de Vibração da Pele do Pescoço no tempo e seu espectro na frequência..



Fonte: Autoria própria.

Figura 3 – Detecção da frequência fundamental (*pitch*).


Fonte: Autoria própria.

$$Erro = \left| \frac{freq. \text{ algortimo} - freq. \text{ PRAAT}}{freq. \text{ PRAAT}} \right| \times 100 \% \quad (1)$$

3. Resultados e discussão

Ao realizar o cálculo de erro relativo médio percentual do método de extração da frequência fundamental baseado na FFT e m relação aos valores fornecidos pelo PRAAT[®], foi obtido um erro de 0,82%. Portanto, empregando esse critério o algoritmo implementafo mostrou-se eficaz.

5. Considerações finais

O trabalhos desenvolvido indica que a técnica para detecção da frequência fundamental de um sinal de VPP, embarcada na placa ESP32 tem potencial para ser usada no monitoramento da saúde vocal, uma vez que os valores obtidos foram próximos aos fornecidos pelo PRAAT[®], software amplamente usado pelos profissionais especialistas em voz. Em trabalhos futuros pode-se implementar um módulo IoT para a transmissão dos valores extraídos, bem como podem ser implementados a extração de novos parâmetros no sinal, como o *jitter* e *shimmer*.

Agradecimentos

Ao IFPB pelo suporte financeiro.

Referências

- BEHLAU, M. **Voz: o livro do especialista**. 1. Rio de Janeiro: Revinter, 2001.
- CHWALEK, P. C.; MEHTA, D. D.; WELSH, B.; WOOTEN, C.; BYRD, K.; FROELICH, E.; MAUER, D.; LACIRIGNOLA, J.; QUATIERI, T. F.; BRATTAIN, L. J. Lightweight, on- body, wireless system for ambulatory voice and ambient noise monitoring. *In: 2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. p. 205-209, 2018.
- PRUKKANON, N.; CHAMNONGTHAI, K.; MIYANAGA, Y. e HIGUCHI, K. VT-AMDF, a pitch detection algorithm. *In: 2009 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 2009*. PRESS, W. H.; TEUKOLSKY, S. A.; VETTERLING, W. T.; FLANNERY, B. P.; Numerical recipes: the art of scientific computing. 3ª ed. New York: Cambridge University Press, 2007.
- RABINER, L.; CHENG, M.; ROSENBERG, A. e MCGONEGAL, C. A comparative performance study of several pitch detection algorithms. *In: IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 5, pp. 399-418, 1976.
- SILVA, I. F.; LIMA, F. S.; GUIMARAES, A. M.; CORREIA, S. E. N.; COSTA, S. L. Proposta de um sistema embarcado para análise da saúde vocal pela vibração da pele do pescoço. *In: XII Simpósio De Engenharia Biomédica*, 2019.