

# Monitoramento da Saúde Vocal empregando Sensor Piezoelétrico e IoT

Lucas Souza, Igor Silva, Thyago Lima e Suzete Correia

**Resumo**—A frequência fundamental é um parâmetro usado na clínica médica para indicar fadiga vocal. Neste trabalho, foi proposto um algoritmo para a detecção da frequência fundamental, no domínio da frequência. Os sinais empregados são da vibração da pele do pescoço, coletados por um colar que emprega um sensor piezoelétrico. Um módulo de transmissão é implementado, empregando o microcontrolador ESP32 e o site Adafruit IO, para permitir o monitoramento da saúde das pregas vocais em tempo real.

**Palavras-Chave**— *Internet das Coisas, Sistemas embarcados, Análise Acústica, Tecnologias vestíveis.*

## I. INTRODUÇÃO

Tecnologias vestíveis permitem o monitoramento contínuo de atividades físicas e de dados biológicos, fisiológicos ou comportamentais de seus usuários. Para a coleta de dados são usados sensores, que podem ser acoplados a itens de vestuário como sapatos, acessórios, roupas, luvas e *smartphones*, ou podem ser anexáveis a pele, como adesivos [1].

Os dispositivos vestíveis têm revolucionado o monitoramento remoto, tornando-o mais eficiente, conveniente e amigável ao paciente. De acordo com Quian e Long [2], os dispositivos vestíveis apresentam vantagens como: mobilidade sem fio, interatividade, inteligência, sustentabilidade, durabilidade, miniaturização, usabilidade e portabilidade. Em aplicações médicas, os dispositivos vestíveis têm sido utilizados para o monitoramento de glicose no sangue [3], monitoramento dos batimentos cardíacos em indivíduos com doenças cardiovasculares [4], detecção pré-sintomática da COVID-19 [5] e a medição da função de deglutição [6].

Profissionais especialistas em voz, como otorrinolaringologistas e fonoaudiólogos, utilizam diversos métodos para avaliar e classificar os distúrbios vocais. A avaliação perceptivo-auditiva e a videolaringoscopia são considerados padrão-ouro, por fornecerem informações anatômicas e funcionais da laringe. Entretanto, no caso de abuso vocal, tais exames não conseguem determinar o que desencadeou a desordem vocal [7]. Nesse caso, o autorrelato do paciente, sobre a utilização de sua voz no cotidiano, pode ser considerado, mas não é confiável, visto que, pode ocorrer a omissão ou esquecimento de determinados hábitos relacionados ao uso da voz [8].

Holbrook, Rolnick e Bailey [9] demonstram que o monitoramento contínuo das atividades de voz e a utilização de um *biofeedback* (sinalizando para o paciente quando ocorre o abuso vocal), é uma ferramenta extremamente importante para detecção das causas de distúrbios relacionados à voz e a correção de hábitos vocais. Diversos esforços têm sido feitos

para o desenvolvimento de dispositivos vestíveis que permitam a avaliação contínua das atividades vocais de um indivíduo. Dentre as tecnologias existentes, estão presentes o uso de sensores para a análise das Vibrações da Pele do Pescoço (VPP), como acelerômetros, piezoresistivo e piezoelétrico, e microfones para a detecção das atividades de voz [8 -11].

Através do monitoramento da VPP é possível extrair parâmetros que podem quantificar a eficiência e a qualidade do sistema vocal de um locutor. Um desses parâmetros é a frequência fundamental (*pitch*) do sinal de voz, que reflete o número de ciclos vibratórios das pregas vocais por segundo. Essa é uma medida essencial, pois a elevação do *pitch* acima do seu nível normal é um indicativo que o indivíduo entrou em fadiga vocal [12]. Estudos apontam que esse tipo de análise é promissora, por ser uma técnica não invasiva, causar o mínimo de desconforto ao paciente, ter robustez a ruídos ambientais e possibilitar uma melhor detecção dos fonemas sonoros, por estarem sendo extraídos próximo à fonte glótica [13].

Mehta et al. [10] propuseram o uso do *Ambulatory Phonation Monitor* – APM (Monitor de fonação ambulatória), utilizando como sensor um acelerômetro em miniatura colocado no pescoço, logo acima do entalhe esternal, para medir as vibrações induzidas na pele. O dispositivo possibilita estimar as medidas de frequência fundamental, nível de pressão sonora e tempo de fonação por até 14 horas diárias. Este dispositivo permite uma interface com *smarthphone*. Com este estudo foi constatado que para notar a variação da pressão sonora e frequência fundamental, é necessário o monitoramento de 20 horas continuamente.

Chwalek et al. [11] propuseram um dispositivo vestível compacto, leve 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 no início do osso esterno e um microfone para detecção do nível de pressão sonora. O protótipo consiste em um colar, que aloca todos os dispositivos como acelerômetro, microfone, bateria e dispositivo para conexão sem fio. Com este estudo, foi possível detectar a frequência fundamental, o nível de pressão sonora e o tempo de fonação.

Silva et al. [8] propuseram um colar para capturar os sinais de vibração da pele do pescoço empregando um sensor piezoelétrico de 35 mm de diâmetro, situado no início do osso esterno no pescoço. Foi realizada uma comparação com a análise acústica para 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. Também foi possível extrair medidas de *jitter*, *shimmer* e HNR (*Harmonics-to-Noise Ratio*) com o dispositivo. O colar possui baixo custo de produção, é confortável e adaptável a qualquer medida de circunferência de pescoço. No entanto, as medidas foram extraídas empregando o software PRAAT©, amplamente usado em análise e síntese da

fala, não tendo, esse estudo, desenvolvido um sistema embarcado com essa finalidade.

Neste trabalho é apresentado o desenvolvimento de um módulo de transmissão, empregando Internet das Coisas (IoT – *Internet of Things*), para ser incorporado na tecnologia vestível desenvolvida por Silva et al. [8]. Para tanto foi embarcado um algoritmo de detecção de *pitch*, baseado na FFT (*Fast Fourier Transform*), em uma placa ESP32. Como plataforma para visualização da frequência fundamental, foi utilizado o site Adafruit IO, uma solução gratuita e simples, onde é possível enviar dados de um dispositivo IoT e apresentá-los em um gráfico, que se atualiza em tempo real. Com essas características é possível monitorar as variações na frequência fundamental de um indivíduo e manter um histórico para uma análise posterior de um profissional da saúde.

## II. MATERIAIS E MÉTODOS

Os sinais de vibração da pele do pescoço foram obtidos através do colar proposto por Silva et al. [8] e fazem parte de uma base de dados aprovada pelo Comitê de Ética em Pesquisa (CEP), do IFPB, de nº 3.577.570. Essa base possui 128 sinais de VPP, de 32 voluntárias, onde cada uma participou de quatro sessões de atividade de fala para a aquisição dos dados. 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.

Os sinais de vibração foram gravados em um arquivo e apresentados ao microcontrolador ESP32, para a detecção da frequência fundamental e seu envio, via comunicação por Wi-Fi, permitindo o monitoramento em tempo real, em uma página da Internet. A Fig. 1 ilustra as etapas envolvidas na metodologia.

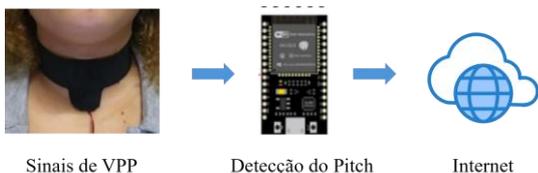


Fig. 1. Metodologia empregada para o monitoramento da frequência fundamental dos sinais de VPP.

### A. Detecção da frequência fundamental

A frequência fundamental corresponde a frequência mais baixa presente no espectro de um sinal periódico, identificada como sendo o primeiro pico de frequência. A Fig 2. apresenta a representação de um sinal de VPP, no domínio do tempo e da frequência.

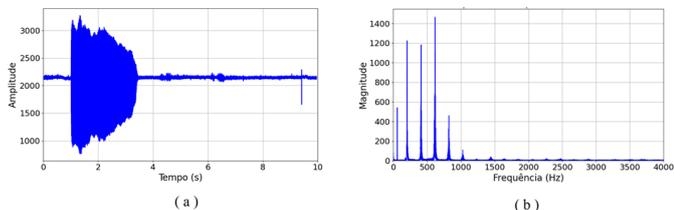


Fig. 2. Sinal de VPP: (a) Domínio do tempo; (b) Domínio da frequência.

Na literatura podem ser encontrados diversos algoritmos capazes de detectar a frequência fundamental de sinais de áudio. Há 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 FFT, análise cepstral e máxima verossimilhança [14-16].

A detecção da frequência fundamental ou *pitch*, foi implementada na linguagem de programação C e embarcada na plataforma de desenvolvimento ESP32. Para a identificação do pico que corresponde ao *pitch*, foi desenvolvido um método, para o domínio da frequência, que é uma adaptação, dos algoritmos apresentados por Prukkanon et al. [14] e Quin et al. [15], que empregam a técnica no domínio do tempo AMDF. A ideia foi ajustar o processo de busca de todos os vales contidos no algoritmo AMDF, a ordenação e seleção do menor vale, que corresponde a frequência fundamental, para determinar os picos da FFT e determinar qual o que ocorre na menor frequência, obtendo-se assim a frequência fundamental.

Para tanto, inicialmente foi definida uma faixa de frequência de interesse no espectro. A faixa utilizada durante os experimentos foi de 100 Hz a 900 Hz. A magnitude e a frequência de cada ponto, dentro dessa faixa, foram armazenadas em uma estrutura de dados, como sendo um par ordenado (x,y). Em seguida, 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.

O primeiro elemento da estrutura de dados ordenada é considerado como sendo o pico correspondente à frequência fundamental. Esse elemento pode ser de fato o *pitch* do sinal ou pode ser, provavelmente, um outro harmônico que possui uma magnitude elevada (como pode ser visto na Fig. 2(b)). Na sequência, 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, bem como a distância entre o pico atual e esse elemento seja maior ou igual a um valor 100 Hz (determinado empiricamente), então esse elemento é guardado como o novo pico atual. Os elementos da estrutura são percorridos até se alcançar um limite inferior de magnitude, como pode ser visto na Fig. 3. Nos experimentos esse limite inferior foi definido como sendo 20% da magnitude máxima encontrada na estrutura de dados (ou seja, do primeiro elemento da estrutura). Esse processo de detecção da frequência fundamental pode ser observado na Fig. 3.

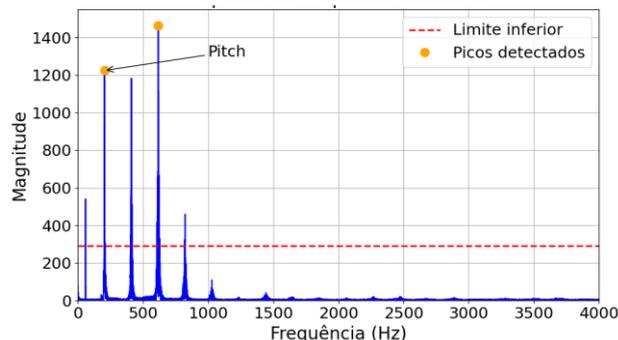


Fig 3. Detecção do pitch de um sinal de VPP.

B. Monitoramento vocal

Por conta da conectividade Wi-Fi, a ESP32 foi escolhida para transmitir a frequência fundamental do sinal de VPP, utilizando o *framework* ESP-IDF da empresa Espressif.

O site Adafruit IO foi empregado para a visualização em tempo real do parâmetro transmitido. Essa plataforma é atualizada em tempo real, à medida que novos valores chegam. Esse site armazena durante 30 dias os dados enviados e permite que eles sejam posteriormente baixados.

A Adafruit IO fornece uma API (*Application Programming Interface*) para que a comunicação com o dispositivo IoT seja realizada. Neste trabalho foi utilizada a API baseada no protocolo MQTT (*Message Queuing Telemetry Transport*), um protocolo ideal para aplicações de IoT por ser leve e simples de ser utilizado.

III. RESULTADOS

Para avaliar o algoritmo de detecção da frequência fundamental proposto, foi feito um comparativo entre os resultados obtidos pelo software PRAAT© com os obtidos utilizando-se os algoritmos da AMDF [14] e da FFT modificado. Esse comparativo foi quantificado através do cálculo do erro relativo percentual, dado na Equação (1). Ao realizar o cálculo do erro relativo percentual dos métodos de extração do *pitch*, foi obtido um erro de 5,32% para o método que usa o AMDF e 0,82% para o método que usa a FFT, indicando a eficácia do método embarcado.

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

Na Fig. 4 é apresentado o gráfico fornecido pelo site Adafruit IO, contendo a frequência fundamental de um sinal de VPP, registrada ao longo do tempo. Esse gráfico fica disponibilizado para acesso remotamente, permitindo o monitoramento da saúde das pregas vocais.

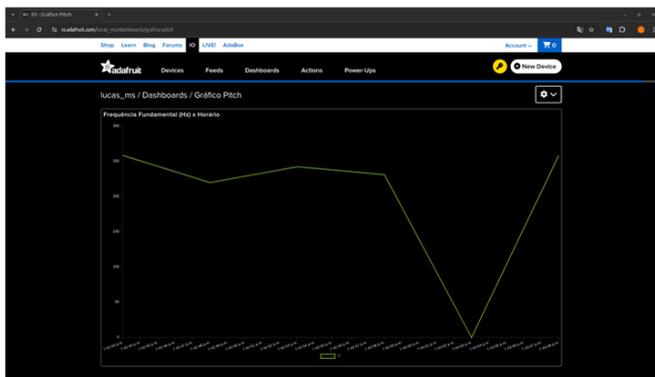


Fig 4. Gráfico da frequência fundamental de um indivíduo ao longo do tempo no site Adafruit IO.

IV. CONCLUSÕES

Foi apresentado um sistema para o monitoramento de valores da frequência fundamental da voz, obtidos por meio de sinais da vibração da pele do pescoço. Um algoritmo para a detecção do *pitch*, baseado na FFT foi proposto e embarcado na plataforma ESP32. Os valores obtidos foram bem próximos

dos fornecidos pelo software comercial PRAAT©. A visualização e acompanhamento dos valores da frequência fundamental, foi feita no site Adafruit IO.

Em trabalhos futuros pode-se testar diferentes plataformas para visualização dos parâmetros transmitidos que tragam uma interface mais personalizável e com mais informações. Outros avanços que podem ser implementados é a extração de parâmetros no sinal, como o *jitter* e *shimmer*, que também devem estar embarcados na própria placa ESP32.

AGRADECIMENTOS

Ao Instituto Federal da Paraíba pelo suporte financeiro.

REFERÊNCIAS

- [1] V. Linh, S. Han, E. Koh, S. Kim, H. Jung, J. Koo, "Advances in wearable electronics for monitoring human organs: Bridging external and internal health assessments", *Biomaterials*, v. 314, 2025.
- [2] RC. Qian, YT. Long, "Cover Feature: Wearable Chemosensors: A Review of Recent Progress", *ChemistryOpen*, v. 2, Jan 2018 Jan 18.
- [3] A. K. M and S. Baskar, "Internet of Things based Real-Time Hyperglycemia and Hypoglycemia Monitoring using Wearable Biosensors", 2022 International Conference on Electronics and Renewable Systems (ICEARS), 2022.
- [4] G. Kaspar, K. Sanam, G. Gholkar, NR. Bianco, S. Szymkiewicz , D. Shah, "Long-term use of the wearable cardioverter defibrillator in patients with explanted ICD ". *Int J Cardiol*. Dez 2018.
- [5] T. Mishra, M. Wang, A. Metwally, G. K. Bogu, A. W. Brooks, A. Bahmani,... and M. P. Snyder, "Pre-symptomatic detection of COVID-19 from smartwatch data," *Nature biomedical engineering*, v. 4, 2020.
- [6] WY. Shieh, CM. Wang, HK. Cheng, CH. Wang, "Using Wearable and Non-Invasive Sensors to Measure Swallowing Function: Detection, Verification, and Clinical Application", *Sensors (Basel)*, Jun 2019.
- [7] M. BEHLAU, *Voz: o livro do especialista*, Revinter, 2001.
- [8] I. F. Silva, F.S. Lima, A.M. Guimaraes, S.E.N. Correia, S.L. Costa, "Proposta de um sistema embarcado para análise da saúde vocal pela vibração da pele do pescoço", *XII Simpósio De Engenharia Biomédica*, 2019.
- [9] A. Holbrook, M.I. Rolnick, C.W. Bailey CW, "Treatment of vocal abuse disorders using a vocal intensity controller", *J Speech Hear Disord.*, Ago 1974.
- [10] D.D. Mehta, P.C. Chwalek., T.F. Quatieri, L.J. Brattain, "Wireless Neck-Surface Accelerometer and Microphone on Flex Circuit with Application to Noise-Robust Monitoring of Lombard Speech", *Proc. Interspeech* 2017.
- [11] P. C. Chwalek et al., "Lightweight, on-body, wireless system for ambulatory voice and ambient noise monitoring," *2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2018.
- [12] I. Titze. *Princípios da produção vocal*. Salt Lake City: National Center of Voice and Speech, 2013.
- [13] A. G. Alzamendi, "Modelado Estocástico de la Fonación y Señales Biomédicas Relacionadas: Métodos en Espacio de Estados Aplicados al Análisis Estructural, al Modelado de la Fonación y al Filtrado Inverso", Dissertação (Mestrado em Engenharia Elétrica). Universidad Nacional Del Litoral, Santa Fé, Argentina, 2016.
- [14] N. Prukkanon, K. Chamngongthai, Y. Miyanaga, K. Higuchi, "VT-AMDF, a pitch detection algorithm," *2009 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, 2009.
- [15] L. Qin, Q. Li and X. Guan, "Pitch extraction for musical signals with modified AMDF," *2011 International Conference on Multimedia Technology*, 2011.
- [16] L. Rabiner, M. Cheng, A. Rosenberg e C. Mcgonegal. "A comparative performance study of several pitch detection algorithms", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 5, pp. 399-418, 1976.