Design and Analysis of Noninvasive Photoglottography Device

Vaughan Altmann, Deepta Bharadwaj, Ivy Goodwin, Isabelle Villegas, and Erica Yang

*Shu Chien-Gene Lay Department of Bioengineering, University Of California, San Diego

***Abstract*— This study presents the development of a non-invasive photoglottography device for assessing the behavior of the glottis during speech and breathing. Traditional methods employing invasive procedures limit the study of laryngeal abnormalities due to their discomfort and complexity. Our device utilizes an external infrared LED array paired with an optical sensor, circumventing the need for invasive endoscopes. The signal processing circuitry, including a sensor amplifier, bandpass filter, and non-inverting amplifier, facilitates the capture and analysis of vocal fold movements. Through experimental validation and analysis, we demonstrate the device's efficacy in detecting glottal states and capturing vocal cord vibrations. Design improvements, such as incorporating additional light sources and enhancing signal processing, are discussed to refine measurement accuracy. The potential applications of this non-invasive photoglottography setup extend to phonetics research, vocal assessment, and monitoring speech disorders, offering new avenues for understanding and addressing laryngeal pathologies.**

***Keywords*— photoglottography, vocal cords, glottis, NIRs, speech disorders**

I. Introduction

The glottis is the part of the larynx that contains the vocal folds, and refers to the space between the vocal cords while they are either in use or disuse. The three basic states of glottal passage are open, such as when someone is breathing but not speaking, in vibration, meaning when someone is speaking, narrowed, in which someone is whispering, and closed, when someone is not speaking nor breathing [3]. Given that the state of the glottis is directly tied to the production of sound, abnormalities in the larynx and/or glottis will result in abnormalities in speech. To study the physical mechanism of speech, photoglottography is the method employed.

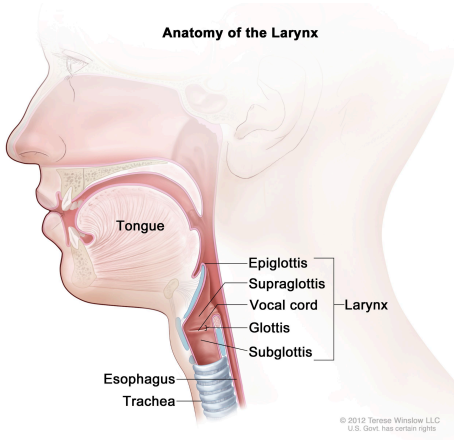

Fig. 1: Anatomy of the larynx [8].

Photoglottography is the technique in which a light source is positioned either above or below the glottis while a light sensor is placed opposite to the light such that the vocal folds are between the light and the sensor. This allows the sensor to capture the position of the vocal folds based on the received light intensity and convert that intensity into a voltage. The prevailing method of photoglottography is highly invasive and involves using a flexible endoscope or fiberscope, inserted nasally, as a light source, while an optical sensor is placed externally on the neck [4]. The predominant issues with this method are that the use of the fiberscope necessitates the presence of a medical specialist and the view of the glottis can sometimes be obstructed during certain speech patterns, like with the formation of back vowels in words like "rule" or "pole" [4]. If the glottis is obstructed from the light source, there will be no light intensity for the source to determine a reading from. An additional consequence of this method being invasive is that the study of the larynx/glottis has been limited due to the reluctance of researchers to subject people to such an invasive procedure [5]. Therefore, the goal is to design a *noninvasive* photoglottography device that is able to measure the aperture of vocal chords, possibly for phonetics research, or to aid in the identification and study of speech disorders [5].

II. Methods

A. Overview

The general device schematic for a noninvasive photoglottography device includes using an external infrared, or near infrared, LED array as an alternative light source to an internal fiberscope or flexible endoscope [7]. This is because infrared light can penetrate up to 5 mm into the skin while the trachea is, on average, about 3 mm thick [4]. Visible light spectra, while also able to penetrate skin, is more susceptible to absorption. The chosen light source to work in conjunction with the IR LED array is BPW 34 because it is compact and one of the smaller and cheaper sensors. The attached signal processing circuitry is designed to allow for the determination of how fast and how open the vocal cords are during speech and breathing. Once these signals have been captured and processed, they would be sent to some additional digital recording device for visual analysis.

B. Full Circuit

The primary focus of this device design will be on the signal processing component, the input of which is from the optical sensor and the output of which is relayed to a recording device for graphical analysis. The three main components of the circuit are an amplified sensor, a bandpass filter, and a non-inverting amplifier. Assume

that all op amps are ideal and have a power supply of ± 3.3V.

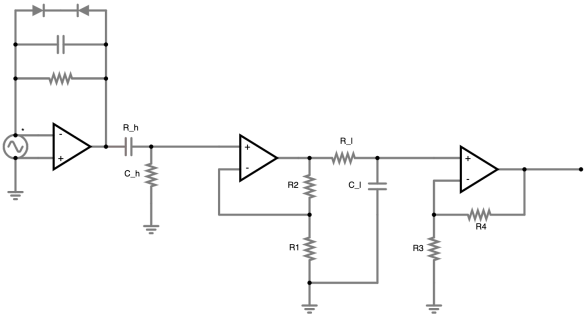

Fig. 2: This figure displays the circuit for the signal processing component. It consists of a sensor amplifier, a band pass filter, and a noninverting amplifier, though the sensor amplifier is not pictured for simplicity [9].

III. Results And Analysis

A. Sensor Amplifier

In the circuit diagram pictured below, the voltage source represents the optic transducer, which in this case is BPW34. The average current produced by the photosensor is expected to be 7 nA. The current is driven through the resistor by the op amp creating an output voltage of 0.42 V, essentially creating a 0.42V AC voltage source. The capacitor is in parallel to the 60MΩ resistor to reduce noise. The Zener diodes have a breakdown voltage of 2V, and are present to prevent op amp saturation due to artifacts/uneven signals.

$$V = IR = 0.42V$$

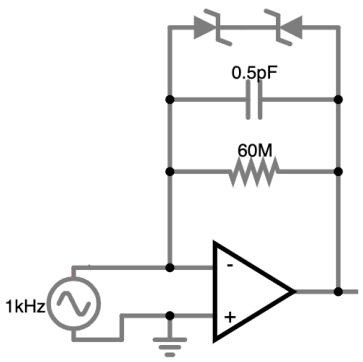

Fig. 3: Circuit diagram for sensor amplifier component [9].

B. Band Pass Filter

Next, the Bandpass filter has two op amp circuits; The first op amp circuit is a high pass filter with a cutoff frequency of 90 Hz while the second op amp circuit is a low pass filter with a cut off frequency of 2000 Hz. These cutoff frequencies reflect how 90 Hz is the low end of the male vocal cord frequency range (also removes 60 cycle frequency due to AC source) and 1000 Hz is the high end of the female vocal cord frequency range. Most human speech lies between 110 and 200 Hz frequencies, which is ideal for our filter.

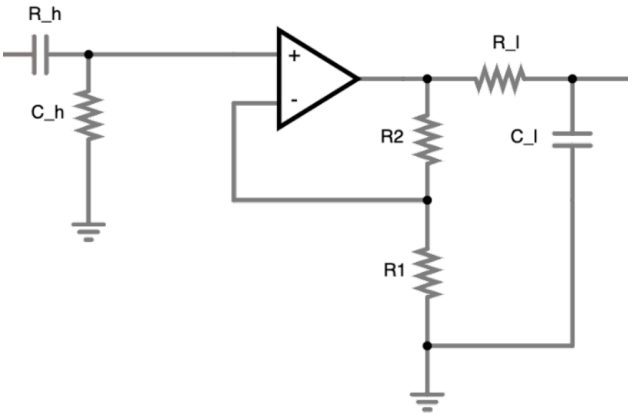

Fig. 4: Circuit diagram for band pass filter [9].

a) Bode Plot of Band Pass Filter:

Bandpass filter attenuates incoming frequencies. Due to the three high pass filters, in the first part of the filter, there is a sharp peak at the 90 Hz cutoff frequency. The lower cutoff frequency is 90 Hz and the higher cutoff frequency is 1000 Hz. The resonant frequency is 300 Hz.

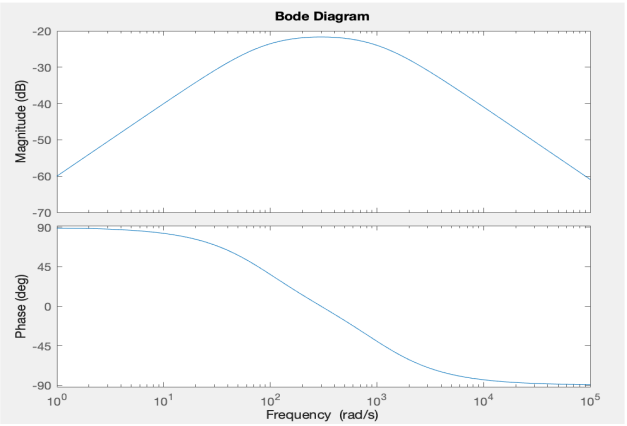

Fig. 5: Bode plot of band pass filter [9].

b) Output Voltage:

Max output voltage at resonant frequency 300 Hz was found to be approximately 2.3V as seen in Fig. 5. The max output voltage at average speaking frequency of males and females ~155 Hz was determined to be approximately 2.1V, max voltage output for males shown in Fig. 6.

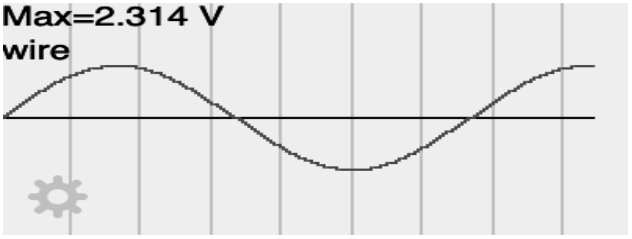

Fig. 6: Max output voltage at resonant frequency 300 Hz [9]

Fig. 7: Max output voltage at average speaking frequency of males ~ 110 Hz [9]. Max output voltage at average speaking frequency of females ~ 200 Hz (not shown) was found to be 2.259V [9].

c) Equations used to determine component values:

For the band pass filter:

$$f_{cut-off} = \frac{1}{2\pi RC}$$

High pass filter with $f_{cut-off}$ = 90Hz; C_h = 1nF:

$$R\_h = \frac{1}{2\pi C\_h f_{cut-off}} = 1.8M\Omega$$

Low pass filter with $f_{cut-off}$ = 1000Hz; C_l = 1nF:

$$R\_l = \frac{1}{2\pi C\_l f_{cut-off}} = 160k\Omega$$

The 200kΩ and the 100kΩ have been added to amplify the signal to be able to check in advance.

$$Gain_{bp} = 1 + \frac{R2}{R1} = 1 + \frac{200k\Omega}{100k\Omega} = 3$$

C. Non-Inverting Amplifier

This amplifier amplifies the signal by 2 times the original voltage, and brings the signal range around +1.3V to -1.3V. Amplifier and bandpass filter brings all signals within the human vocal cord range to above 1V, while weakening those outside of it. The open wire goes to computer/graphing programs.

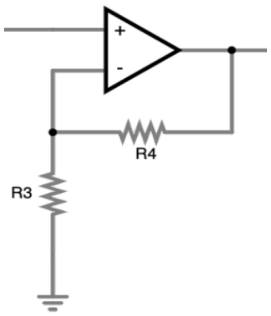

Fig. 8: Circuit diagram for non-inverting amplifier [9].

d) Equations used to determine component values:

$$Gain_{Amp} = 1 + \frac{R4}{R3} = 1 + \frac{100k\Omega}{100k\Omega} = 2$$

Which gives us a theoretical total gain of:

$$Gain_{bp} \times Gain_{Amp} = 6$$

$$V_{out} = V_{in} \times Gain_{bp} \times Gain_{Amp} = 2.52V$$

At the resonance frequency the $Gain_{bp}$ is 3 because we have a 3db loss at the resonance frequency.

III. Discussion

A. Design Improvements

For potential design improvements, it must be considered how, in practical settings, the light variation is still extremely faint due to tracheal absorption, meaning the sensor has trouble with even slight anatomical variations of the glottal position [2]. The effect of tracheal absorption may also read poorly and be highly susceptible to artifacting, interference, or false negatives. Multiple light sources or more sensitive sensors with a larger gain ratio may improve measurement accuracy. Moreover, pairing the device with a microphone may further improve the pairing of sound to speech to facilitate phonetics research. Additionally, based on the bode plot, the band pass filter lacks strong attenuation, meaning the signal refractions that introduce noise below 90 Hz or above 1000 Hz will not be adequately filtered out.

B. Design Applications

Some possible applications of this design are that this particular photoglottography setup could encourage/facilitate the exploration of experimental phonetics, vocal assessment, the phonatory effect of diseases, disorders of the larynx, and/or the development of wearable systems for monitoring speech.

Acknowledgments

Thank you to Professor Cauwenberghs and our wonderful TAs Adyant Balaji, Vikrant Jaltare, and Samira Sebt for their help throughout this quarter and feedback for this project.

References

[1] Amelot, A., Sathiyanarayanan, D., Maeda, S., Honda, K., & Crevier-Buchman, L. (2020). Validation of a Noninvasive System to Observe Glottal Opening and Closing: External Photoglottography (ePGG). *11th International Conference on Voice Physiology and Biomechanics*. East Lansing, MI: HAL. https://shs.hal.science/halshs-02510216/document

[2] Chi, Yujie & Honda, Kiyoshi & Wei, Jianguo. (2021). Portable Photoglottography for Monitoring Vocal Fold Vibrations in Speech Production. 6438-6442. 10.1109/ICASSP39728.2021.9413770. http://cic.tju.edu.cn/faculty/Honda/pdf/Yujie_Chi_ICASSP2021.pdf

[3] Esling, John H., & Harris, Jimmy G. (2005). States of the glottis: An articulatory phonetic model based on laryngoscopic observations. In W. J. Hardcastle & J. Beck (Eds.), *A Figure of Speech: A Festschrift for John Laver* (pp. 347-383). Mahwah, NJ: Lawrence Erlbaum Associates.https://www.sciencedirect.com/science/article/pii/B0080448542050987

[4] Furlow PW, Mathisen DJ. Surgical anatomy of the trachea. Ann Cardiothorac Surg. 2018 Mar;7(2):255-260. doi: 10.21037/acs.2018.03.01. PMID: 29707503; PMCID: PMC5900092. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5900092/

[5] Gerratt, B.R., Hanson, D., Berke, G.S., & Precoda, K. (1991). Photoglottography: A clinical synopsis. *Journal of Voice, 5*, 98-105. https://www.sciencedirect.com/science/article/pii/S0892199705801730

[6] NCI Dictionary of Cancer terms. National Cancer Institute. (n.d.). https://www.cancer.gov/publications/dictionaries/cancer-terms/def/glottis

[7] Padalkar MV, Pleshko N. Wavelength-dependent penetration depth of near infrared radiation into cartilage. Analyst. 2015 Apr 7;140(7):2093-100. doi: 10.1039/c4an01987c. PMID: 25630381; PMCID:PMC4418489. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4418489/

[8] Winslow, T. (2017). *Larynx Anatomy*. NIH cancer institute. Retrieved May 21, 2021, from https://visualsonline.cancer.gov/details.cfm?imageid

[9] All analysis and drawings done in falstad and MATLAB.