Speech Detector of Male to Female Using Cepstral Analysis

Engr. Ranil M. Montaril, John Francis T. Olivo, Ian Jay Briñosa, Ariel S. Conde, Angelo F. Fiestada, Shaira Anne Marie P. Rudas, Paul Vincent Sablan and Brian Daniel L. Tiongco

Abstract— Speech is an essential part of communication and requires the use of voice. Nowadays there are systems made that can detect and differentiate male and female voice. A Speech detector of male to female is an application that can enhance the speaker identification and human-machine interactions. Some studies analyzes the speaker by relating to (a) physiological, (b) phonetic, (c) the difference between the quality of their voices, (d) the anatomical difference that arise in their puberty and even sociology and even philosophy. This paper shows the use of Cepstral Analysis in Speech Detector of Male to Female voices by recording a voice or speech using MATLAB and exported in excel. Using the step by step button control of the cepstrum or formant we identify the gender of the speaker and shows the detected wave form.

Index Terms—Speech, Cepstrum Analysis, Speech Detector, MATLAB, Pitch Detection, Voices

— — — — — — — — ◆ — — — — — — — —

1 BACKGROUND OF THE STUDY

Speech is a natural form of communication which requires the use of voice. It is a power that allows us to express our thoughts and feelings by eloquent sounds. In linguistics, speech is a system of communication that uses spoken words [1]. Speech is produced when the air pressure generated by lungs reaches the vocal cords. Then speech begins to resonate in the nasal cavities according to the position of lips, tongue and other organs in the mouth [2].

Speech Detector of a male and female is an important application that can improve speaker identification and segmentation and even human-machine interactions. Since the advancement in the field of digital signal processing, it is more likely that systems which uses a voice interaction software will take the place of the keyboards in the future.

A gender speech detector is a complicated task since the differences between a male's voice and a female's voice are linked to complex and multidiplicinary issues for they not only refer to its fundamental and resonant frequencies and perceptual measurements but also to its differences in their vocal organs which are anatomy, physiology, sociology and even philosophy [3].

The human voice is consist of a sound made using the vocal folds. In order to analyze the gender, previous work has studied the variations between male and female speech.

These includes the (a) physiological, (b) phonetic and (c) the differences between their voice qualities [2]. A study also states that a majority of authors, the variations can only be consider by the anatomical and the physiological differences that arise during puberty [3]. A male's voice tends to vibrate more slowly than the female's since their vocal folds becomes thicker and longer. Longer wavelength makes lower pitch.

Various papers and researches have been carried out on the aim to improve gender and speech detection. Generally, there are three primary approach in the implementation and construction of a speech detector of male to female voice. (a) The first approach uses pitch as a discriminating factor and use the labelled data to identify the speaker's gender. (b) The second approach is concerned with acoustic features and unlabeled data to identify the speaker's gender and (c) the third approach is the combination of pitch models with acoustic models to form a fused model [2].

In this paper, the Speech Detector of Male to Female voice uses Cepstral Analysis to identify it's the gender of the speaker wherein the proponents will record a voice or a speech signal using Matlab and will export the raw data in excel. The program will have a step-by-step button-controlled generation of the cepstrum or formants and will be able to identify the gender of the speaker and show the recorded voice's waveform.

- *I.J. Briñosa, A. Conde, A. Fiestada, J.F. Olivo, S.A.M. Rudas, P.V. Sablan and B.D. Tiongco are students of the Department of Computer Science, College of Computer and Information Sciences at the Polytechnic University of the Philippines, Manila, Philippines.*
  Email: {ianjay.brinosa, asconde1997, angelofiestada, olivo.jfrancis, shaira.rudas.sr, vincentpaul_012, bdtiongco} @gmail.com

- *R.M. Montaril is a professor of the Department of Computer Science*
  Email: rmmontaril@pup.edu.ph

2 PROBLEM STATEMENT

This goal of this project is to determine the gender of speaker's voice in the recorded audio.

Specifically the project aims to answer the following question:

1. What is the degree of accuracy of identifying the gender in voice inputs using Cepstral Analysis?

3 APPLICABLE RELATED STUDIES

Fast Fourier Transform and Inverse Fast Fourier Transform. In the study of [4] used two signal analysis techniques FFT together with LPC. Using these techniques, the speech signal first analyzed and different parameter of speech signal is obtained. The parameters of a word were obtained then these parameters values can be used in many software for word recognition.

Orthogonal frequency division multiplexing (OFDM) is a special case of multicarrier transmission where a single DataStream is transmitted over a number of lower rate subcarriers. In the study of [5] IFFT/FFT are designed for OFDM. Inverse Fast Fourier Transform (IFFT) was used for signal generation at the transmitter which is used to convert frequency domain to time domain and Fast Fourier Transform (FFT) which is used to convert time domain to frequency domain at the receiver. The implementation of an 8 Point IFFT processor is done by using components like adders, subtractors, multipliers and buffers In IFFT the twiddle factor values are of unsigned values which have to be converted into binary form for the multiplication purpose.

Cepstrum and Cepstral Analysis. In the study of [6] they used the cepstral coefficients to extract spectral feature from the VAG signals on order to represent the VAG for digital signal processing. Cepstral analysis to derive a simple linear acoustic model in traditional speech processing. The speech signal is produced by convoluting an excitation waveform with vocal filter then converting the waveform into frequency domain. Fourier Transform was applied onto the vocal tract filter to obtain the cepstrum which is divided into real and complex kepstrum's. Cepstral coefficient is capable of disclosing higher and more prominent discriminant information by weighting on lower band frequencies and exhibit superior separability in feature. The cepstral analysis of VAG signal has been showing great potential for future research given the high performance in intra-subject analysis.

Sheik [7] used the Mel Frequency Cepstral Coefficients (MFCC) to represent the shape of the vocal tract using the short-term power spectrum

thus, trying to approximate human auditory system responses. As the MFCC represents a short-term power spectrum so they are considered to be short term features.

Hon [8] used a cepstral mean subtraction to amplify the perceived human voice in linear scale. The use of The MFCC scale related to perceived frequency is measured frequency, which tells exactly how to space filter banks and how wide they should be made. The features match more closely to what humans hear by incorporating such scale. Then the log Mel spectrum is converted back to the time-domain. The Mel spectrum coefficients are natural numbers, we are using Discrete Cosine Transform. MFCC detection is to split the speech signal into small frames in order to get relatively stable speech signal within each frame the calculation of periodogram for each frame can be estimated by the power spectrum FFT to identify frequencies presented in the frame. Depending on location in the cochlea of the vibrations, different neural excitations inform the brain of certain frequencies.

In the research made by Jlassi et. al [9] used a new method for estimating the pitch from the speech signal which consists of analyzing real cepstrum by the multiscale product (MP) using continuous wavelet transform having one vanishing moment. They framed the voiced signal, then calculated the real cepstrum of each frame. They compute the MP of the cepstrum. The MP is the product of the WTC at three scales. The method was evaluated by the Keele database under clean and noisy conditions. Experimental results indicate that the gross pitch errors are lower than the compared methods under clean and noisy conditions.

Hua et. al [10] used Cepstral analysis in signal processing procedure for audio watermark detection in echo-based audio watermarking systems. It provides rigorous derivations to reveal the advantages of using real cepstrum than complex cepstrum in echo-based audio watermark detection. by noting that both real part and imaginary part cepstral contain a full version of the echo kernel coefficients, which can be appropriately combined to obtain a composite cepstrum to further suppress the interferences. the joint detection scheme over conventional approach using real cepstrum are illustrated. The detection robustness is evaluated using the peak-to-average power ratio. The relationships among echo length, and echo.

In the study of Alexendri et. al [11] revolutionized the speech recognition system by investigating the two-well-known representation of speech signals. We introduce a pseudo-autocorrelation domain, which can be interpreted as a "Root-cepstral domain", and we show how non-parametric cepstral and linear predictive analyses converge to the same optimal solution.

The use of Fourier Transform for cepstrum analysis [12] in seismic signal processing for resolving subsurface structural properties. Homomorphic deconvolution using

cepstrum analysis has been an effective method for wavelet estimation. The inverse of the Fourier transform of the logarithm of a signal's Fourier transform is the cepstral domain representation of that signal. The convolution operation of two signals in the time domain becomes an addition in the cepstral domain. The fractional Fourier transform (FRFT) is the linear chirps whereas the kernel is composed of complex sinusoids.

Hamming Windowing. In the study of [13] they used an advance peak windowing method that overcomes the drawback of the conventional windowing method while maintaining almost the same spectral mask and providing more efficient bit-error-rate performance. Through a numerical analysis and computer simulation, they showed that the scheme can be implemented by using matrix form and exceeds the conditional windowing method. Hamming window is one of the simplest window functions, and a member of the cosine-on-pedestal family.

Formant Estimation. The used of method Deep Formants for estimating and tracking formant frequencies using deep network trained on the aforementioned annotated corpus [14]. In the task of formant estimation, the input is a stationary speech segment such as the middle of a vowel and the goal is to estimate the first 3 formants. In the task of formant tracking the input is a sequence of speech frames and the goal is to predict the sequence of the first 3 formants corresponding to the input sequence. In both tasks the signal is represented using two sets of acoustic features. While in the study of [15] a complex linear prediction analysis was used as a method for estimating the formant frequencies of noisy speech.

Pitch Estimation. During voiced speech segments, the regular glottal excitation of the vocal tract produces energy at the fundamental frequency and its multiples. The changing pitch carries a good part of the auditory message. [13] It discriminates words in tonal languages, allows expressing emotions discriminates questions from statements, and allows emphasizing parts of an utterance.

In the study of [16] they presented and evaluated an algorithm for estimating the fundamental frequency or pitch in speech signals. They utilized models of voiced speech on three levels. The first level is the time frame for spectral analysis. Then they decomposed the short-term spectrum using non-negative matrix factorization (NMF). The spectrum was represented as a weighted sum of harmonic templates and templates for nonharmonic speech. NMF outputs a matrix that holds harmonic energy along possible pitch contours.

4 APPLICABLE EQUATIONS

Fast Fourier Transform and Inverse Fast Fourier Transform. The Fast Fourier transform (FFT) algorithm computes the discrete Fourier transform (DFT) of a sequence, or its inverse (IFFT) [17]. It uses complex exponentials (sinusoids) of various frequencies as its basis functions. It is often written $f(x) \leftrightarrow F(\omega)$, or $F(f(x)) = F(\omega)$ where F is the Fourier transform operator [18].
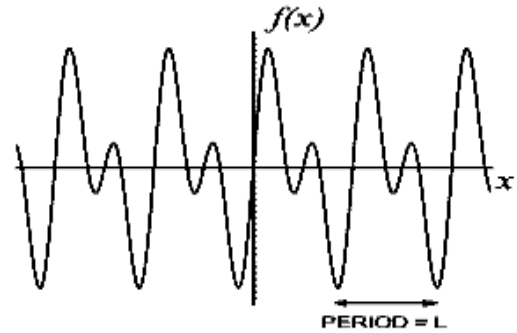
In figure 4.1, A graph of a function $f(x)$ that has period L exhibits the same pattern every L units along the x-axis, so that $f(x + L) = f(x)$ for every value of x.

Fig. 4.1: **Graph of a function** [19]

It explains that if we are familiarized with the complete period of the function, we can complete a greater amount of interval x. That repetition expounds a fundamental spatial frequency $k = \frac{2\pi}{L}$ which may be used in giving the first approximation to its recurrent format $f(x)$:

$$f(x) \simeq c_1 \sin(kx + \alpha_1) = \alpha_1 \cos(kx) + b_1 \sin(kx)$$

where the given symbols with a subscript of 1 are known as constants and determines the amplitude and phase of the first approximation.

The IFFT is as simple as compared to the implementation the DFT and and its 8-point DIF output is derived from the input directly [5]. Figure 2 shows an implementation that can be done in three stages.
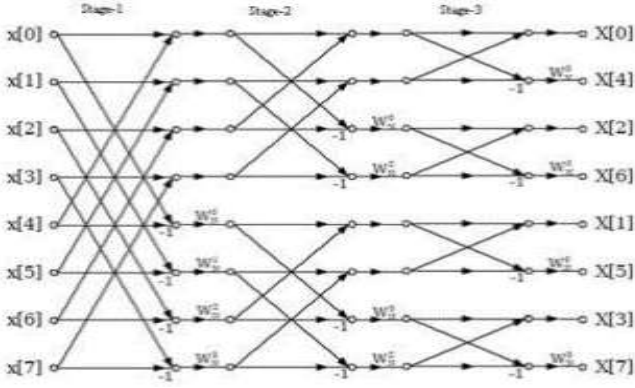
Fig. 4.2: **Stages**

Cepstrum and Cepstral Analysis. Central analysis [20] is based on the observation that $x[n] = x_1[n] * x_2[n] \Leftrightarrow X(z) = X_1(z)X_2(z)$ by taking log of X(z), $\log\{X(z)\} = \log\{X_1(z)\} + \log\{X_2(z)\} = \hat{X}(z)$. The real cepstrum is defined as:

$$c_x[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log|X(e^{jw})|e^{j\omega n}d\omega$$

and the complex cepstrum as:

$$\hat{x}[n] = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log[X(e^{j\omega})]e^{j\omega n}d\omega$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} [\log|X(e^{j\omega})| + j\arg(X(e^{j\omega}))]e^{j\omega n}d\omega$$

where arg() represents the phase.

Hamming Windowing. The generalized hamming window family is formulated by multiplying a rectangular window by one period of a cosine. The benefit of the cosine tampering is the lower side lobes and the price for the benefit is that the width of the main-lobe doubles in width [21]. Its two (2) well-known members are (a) Hann and (b) Hamming Windows.

The generalized hamming window family's basic idea is illustrated in figure 4, where the center dotted waveform is the aliased sinc function [21]:

$$asinc_M(\omega) \triangleq \frac{\sin(\frac{M\omega}{2})}{M \cdot \sin(\frac{\omega}{2})}$$

The aliased sinc function refers to the fact that it may be simply obtained by sampling the length-T continuous-time rectangular window, which has Fourier transform sinc $(f\Gamma) \triangleq \sin(\pi f\Gamma)/(\pi f\Gamma)$ [22].
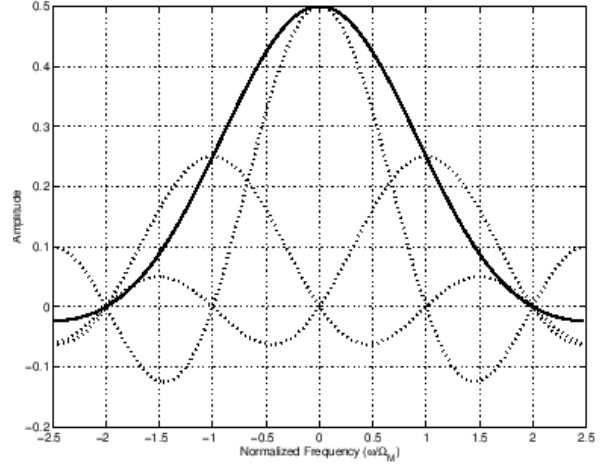
The aliased sinc function approaches to the sinc function as its sampling rate goes to infinity.

$$sinc(x) \triangleq \frac{\sin(\pi x)}{\pi x}$$

The other two (2) dotted waveforms are scaled shifts of the same function, $0.25 \cdot W_R(\omega \pm \boldsymbol{\Omega_M})$. The sum of the three (3) dotted waveforms gives the solid line in figure 4.

Formant Estimation. Formants are the resonance of the vocal tract which have a close relation to the vocal tract geometry [21]. It is linked with peaks in the smoothed power spectrum of speech. The estimation of formant is important in various applications.

Unfortunately, the formant frequencies are onerous to extract from the speech signals. Various research and studies had used methods such as the short-term Fast Fourier Transform, peak-picking on cepstrally smoothed spectra and linear predictive coding (LPC) were used to recognize the location of the formants in a speech signal [21]. The Linear Predictive Coding Model predicts the output of a linear system based on an input x_n and from the previous outputs $S_{n-1}, S_{n-2}, \ldots, S_{n-p}$. It is defined by the mathematical formula:

$$\hat{S}_n = \sum_{k=1}^{p} a_k S_{n-k} + \sum_{k=0}^{N} b_k x_n - k$$

where $\hat{s}_n$ refers to the estimated or the predicted value of s_n. However, the problem is now the identification of the constants a_k and b_k in a way that $\hat{s}_n$ will approximately define the real output s_n as accurately as possible [21]. Various models have been proposed based on the LPC model. The autoregressive model (AR) predicts an output of $\hat{s}_n$ by using the previous outputs and the current input which implies $b_k = 0$ for $k > 0$ [21]. Thus, the problem is to identify only a_k and b_0 which corresponds to an all-pole filter. Mathematically, the Autoregressive model becomes:

$$\hat{S}_n = \sum_{k=1}^{p} a_k S_{n-k}$$

where a_k is known as the predictor coefficient and p is the order of the predictor. Its goal is to identify the parameters a_k so that $\hat{s}_n$ is close to the recorded speech s_n in some frame of the signal [21].

Pitch Estimation. Pitch Estimation is important for recognizing and coding speech [22]. A statistical analysis of the pitch frequency specifies a mean of 123 Hz and a standard deviation of 16 Hz for male voices and a pitch frequency of about as much as twice as the males which is 206 Hz and a standard deviation of 23 Hz for female voices. A pitch estimator must make a:

1. speech or non-speech decision
2. voiced or unvoiced decision
3. estimate of the pitch in the voiced region

Three pitch estimation techniques were proposed to achieve this:

1. Time domain techniques: Peak and valley measurements zero crossings and autocorrelation estimates often with additional post processing logic. Generally these techniques are noise sensitive.
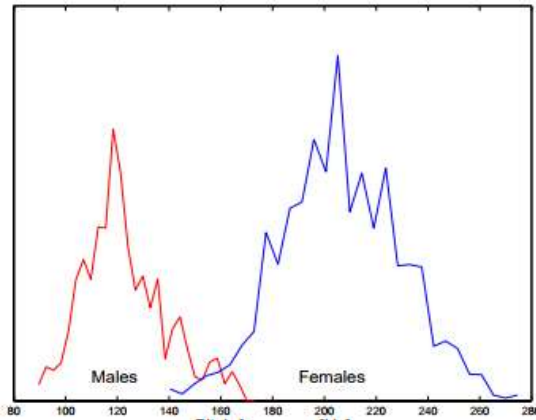

Fig. 4.4: **Histogram of pitch frequency for male and female speakers**

2. Frequency domain technique: Periodicity in the time domain results in useful impulses in the frequency domain at the fundamental and its harmonics. Using short time analysis to extract the frequency information these techniques are sensitive to the length of the analysis window so that their useful dynamic pitch range tends to be limited.
3. Hybrid techniques: The former two techniques combined for instance spectral flattening together with autocorrelation. However this combination does not readily solve the above mentioned weaknesses.

There were four (4) types of error defined when comparing pitch estimates and reference signals of the speech periods where:

$$e = \left| \hat{P} - P \right|$$

Where e is the error, $\hat{P}$ is the estimated pitch period and P is the reference pitch period. The four errors are defined as:

1. If the error is more than 1 ms it is classified as a gross pitch error
2. If the error is less than 1 ms it is classified as a fine pitch error.
3. Misclassification of the transition from the voiced to unvoiced region is classified as a voiced to unvoiced error (V-U error).
4. Misclassification of the transition from the unvoiced to voiced region is classified as an unvoiced to voiced error (U-V error).

Before the estimation of pitch, it is essential to determine if the speech signal is speech and then if speech, if it is voiced [23]. We can use short-time energy and zero-crossing rate. The short time magnitude can be computed every 10 ms for a 10 ms frame length and is mathematically defined as:

$$M(n) = \sum_{m=n-N+1}^{n} |s(m)| w(n-m)$$

where N is the window length and w(m) is a hamming window [22]. Likewise, the zero-crossing rate is computed for each frame [24]. Mathematically, the zero-crossing rate is defined as:

$$Z(n) = \sum_{m=n-N+1}^{n} \left| \left(sgn(s(m)) - sgn(s(m-1)) \right) \right| w(n-m)$$

5 SOFTWARE DESIGN AND ARCHITECTURE

The following shows the models designs and the architecture for the Speech Detector of Male to Female Using Cepstral Analysis.

5.1 System Architecture

The System Architecture of the Speech Detector (Fig. 5.3)

The system starts by initially recording a continuous signal using MATLAB for a maximum of five seconds. After recording, the signal is them saved to an excel file.

The signal data, using a 8KHz sampling rate, frame size of 20ms, and frame step of 10ms, is framed. The result is a signal that is divided into 498 frames, each consisting of 160 samples.

After framing, the each frame undergoes the following process: a) zero padding, b) Fast Fourier Transform, c) Power Spectrum, d) Log of Power Spectrum, e) Inverse Fast Fourier Transform, and f) High-Time Liftering, which are essentially the processes in Cepstral Analysis.

The zero padding is done on every frame to turn each frame into having 256 samples which is a base 2 number.

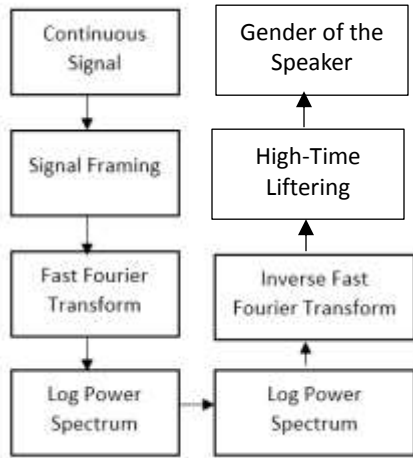(i.e. $log_2 256 = 8$). This is important in the process of Fast Fourier Transform.

Fig. 5.3: **System Architecture**

Then these signals undergo Fast Fourier Transform, specifically a Radix-2 Decimation in Time Fast Fourier Transform. The output is then squared to get the power spectrum. The resulting square outputs are processed to get the logarithm of each. The log of the power spectrum then is processed in an Inverse Fast Fourier Transform. The results of the IFFT are the cepstral coefficients.

The output cepstral coefficients are applied high-time liftering then mode is applied to the results to identify the gender of the speaker.

6   TEST RESULTS

In this paper, the proponents utilized ten (10) audio samples that are of 5 seconds length each. The table on the following page presents the results of the experimentation to identify the speaker's gender.

The frequencies used as basis for identification of the gender are limited to male adults and female adults only which specifically are 85 to 180 Hz and 165 to 255 Hz, respectively.

From the table, there are four columns which are Manual Identification, System Identification, Accuracy, and Error rate.

Test no.	Manual Identification	System Identification
Test 1	Male	Male
Test 2	Male	Unidentified
Test 3	Male	Male
Test 4	Male	Male
Test 5	Male	Male
Test 6	Female	Unidentified
Test 7	Female	Unidentified
Test 8	Female	Female
Test 9	Female	Female
Test 10	Female	Male

7   CONCLUSION

In this paper, we presents a model of Button Controlled Speech Detection for Male and Female Voices using Microsoft Excel and MATLAB for voice recording. By showing the block diagram, state diagram, and code of the system; and is simulated through excel and MATLAB from which we utilized Fast Fourier Transform and Inverse Fast Fourier Transform. We also used the signal framing to minimize the leakage of Cepstrum.

The proponents therefore state that the developed Gender Identification in audio samples using Cepstral Analysis has a 6 out of 10 or 60% accuracy.

8   RECOMMENDATIONS

The proponents of the paper recommend to add voice activity detection module to separate speech segment from non-speech segment and the use of histograms in identifying the gender of the speaker we recommend to a further analysis of the pitch frequency for each consecutive speech segment. We recommend adding such operation to further improve the detection of Male and Female voices.

Moreover, the use of Hamming windowing and / or pre-emphasis technique in preprocessing to smoothen the signal before analysis to show better results.

References

[1]   R. Nordquist, "Speech in Linguistics".

[2]   H. U. Sheikh, *WHO IS SPEAKING? MALE OR FEMALE?*, 2013.

[3]   E. Pepiot, *Voice, speech and gender: male-female acoustic differences and cross-language variation in English and French speakers,* 2013.

[4]   L. Singh, "Speech Signal Analysis using FFT and LPC," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET),* vol. 4, no. 4, pp. 1658-1661, 2015.

[5]   D. R. Bhavani and D. Sudhakar, "International Journal of Science and Engineering Applications," *Design and Implementation of Inverse Fast Fourier Transform for OFDM,* vol. 2, no. 7, 2013.

[6] G. H. Seng, H. Y. Chai and T. T. Swee, "Application of Cepstaral Model on VibroArthrographic Signal fo Future Ostheoarthritis Study," *Recent Advances in Telecommumicatons and Systems.*

[7] H. U. Sheik, "Who is Speaking? Male or Feamale," pp. 33-36, 2013.

[8] Z. Hong, "Speaker Gender Recognition System," pp. 20-26, 2017.

[9] W. Jlassi, A. Bouzid and N. Ellouze, "Pith Estimation Based on the Cepstrum Analysis byj the Multi Scale Product of a Clean and Noisy Speech," *Recent Advances in NonLinear Speech Processing,* pp. 219-225, 2016.

[10] G. Hua, G. Jonathan and V. Thing, "Cepstral Analysis for the Application of Echo Based WatermarkDetection," vol. 10, no. 9, pp. 1850-1861, 20.

[11] P. Alexidre and L. Philip, "Root cepstral analysis: A unified view. Application to speech processing in car noise environments," *Speech Communication,* vol. 12, no. 13, pp. 227-228, 1993.

[12] K. Miah, H. Herera and M. Sacchi, "Application of Fractional Fourier Transform in Cepstrum Analysis," *Recovery, Energy,Environment, Economy,* 2012.

[13] C. A. Bhatnagar, R. Sharma and R. Kumar, "Analysis of Hamming Window Using Advance Windowing Method," *International Journal of Scientific Research Angineering and Technology(IJRSET),* vol. 1, no. 4, pp. 15-20, 2012.

[14] Y. Dissen and J. Keshet, "Formant Estimation and Tracking using Deep Learning".

[15] T. Kaneko and T. Shimamura, "Noise-Reduced Complex LPC Analysis for Formant Estimation of Noisy Speech," *International Journal of Electronics and Electrical Engineering ,* vol. 2, no. 2, pp. 90-94, 2014.

[16] P. a. C. S. Jackson, "Pitch-scaled estimation of simultaneous voiced turbulence-noise in speech.," *IEEE Trans on Speech and Audio Processing,,* vol. 9, no. 7, pp. 713-216, 2001.

[17] "Fast Fourier Transform," 2017 September 30. [Online].

[18] P. Heckbert, "Fourier Transforms and the Fast Fourier Transform (FFT) Algorithm," 1995.

[19] G. S. McDonald, "Fourier Series," 2004.

[20] X. Anguera, "Xavier Anguera," [Online]. Available: http://www.xavieranguera.com/tdp_2011/8-Cepstral-Analysis.pdf.

[21] Formant Estimation using LPC.

[22] S. v. Vuuren, "Pitch Estimation," 1998.

[23] D. G. Childers, M. Hahn and J. N. Larar, "Silent and Voiced/Unvoiced/Mixed Excitation (fourway) Classification of Speech," *Trans. Acoustic, Speech, Signal Processing,* vol. 37, no. 11.

[24] Y. Lau and C. Chan, "Speech Recognition based on Zero Crossing Rate and Energy," *IEEE Trans. Acoustic, Speech, Signal Processing ,* vol. 33, pp. 320-323.

CURRICULUM VITAE

Ian Jay V. Briñosa currently studying at Polytechnic University of the Philippines with the course of Bachelor of Science in Computer Science. He is also a scholar at UNBOUND, a non-government group, His interest are in the field of programming and web development.

Ariel S. Conde is currently a graduating student taking up Bachelor of Science in Computer Science from Polytechnic University of the Philippines. He is interested in the development and applications of Intelligent Agents. He is also proficient in several web and mobile technologies.

Angelo F. Fiestada is currently a graduating student taking up Bachelor of Science in Computer Science from Polytechnic University of the Philippines. Knowledgeable in C, C#, Java, MATLAB`and Web Developement. He is also interested in Artificial Inteligence, Robotics and Data Mining

John Francis T. Olivo is currently taking BSc in Computer Science at the College of Computer and Information Sciences, Polytechnic University of the Philippines, Manila. He graduated from Antipolo National High School on 2014 as 8th Honorable mention. His research interests include Image Processing, Database, and Expert Systems. He is a scholar of the Department of Science and Technology (DOST).

Shaira Anne Marie P. Rudas is currently taking Bachelor of Science in Computer Science at the Polytechnic University of the Philippines. Her research interests includes Web Development, Expert Systems and Image processing. She is also experienced in multitasking and has general skills that are important in working as a part of a team.

Paul Vincent D.L. Sablan is an undergraduate student of Polytechnic University of the Philippines, taking up Bachelor of Science in Computer Science. He has a knowledge on C, C#, Java and HTML and is currently interested in Robotics and Image Processing. He graduated from High School at Cainta Catholic College receiving an award Best in Christian Living and also the treasurer of their class.

Brian Daniel L. Tiongco is a consistent honor student of the Bachelor of Science in Computer Science Major in Research at Polytechnic University of the Philippines who is well-equipped with great analyzing skills and has vast knowledge in programming and computer software applications. He is passionate to explore more of the growing and fast changing technology with a different level of challenges and accomplishments that would further enhance his developing skills.