

DEVELOPMENT OF DIAGNOSTIC PLATFORMS USING SMART TECHNOLOGIES

by

MD. MOTASIM BELLAH

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2015

Copyright © by Md. Motasim Bellah 2015

All Rights Reserved



Dedication

To my family

and

To all enlightened minds who inspired me to pursue this path of knowledge

Acknowledgements

First of all, I would like to thank my family members for staying beside me all the time. I would also like to express my sincere gratitude to my friends, and most importantly my fellow lab mates who offered me countless supports, suggestions, and encouragements throughout my PhD.

I would like to thank my supervisor and mentor Prof. Samir Iqbal who has been always there for me, no matter how difficult the time was. Without his continuous mentorship and support, I couldn't make this far. I would also like to thank all the faculty members from the department of electrical engineering and my honorable committee members, Prof. Donald Butler, Prof. Jung-Chih Chiao, Prof. Nur Yazdani, and Prof. Young-tae Kim for their collaboration and valuable guidance throughout this journey.

The research wouldn't be possible without the financial support from NSF. I would like to convey my heartfelt gratitude towards all the agencies who supported our work financially.

April 9, 2015

Abstract

DEVELOPMENT OF DIAGNOSTIC PLATFORMS USING SMART TECHNOLOGIES

Md. Motasim Bellah, PhD

The University of Texas at Arlington, 2015

Supervising Professor: Samir M Iqbal

Mortality rate of cancer has not significantly dropped over the past six decades because there is no effective tool for early detection. Likewise, mortality rate of cardiovascular (CV) disease remains alarming due to the lack of convenient and accessible tools for monitoring heart conditions regularly over extended periods of time. In addition, children with speech impediments, people with impaired hearing, and learners of new languages do not have an effective means for improving their speech vocalization due to the unavailability of systems that provide quantitative feedback. Moreover, patients with vocal cord disorders have to endure uncomfortable and invasive diagnostic procedures as a result of inadequate non-invasive methods. As a timely diagnosis is the first step towards treating health condition or any problem, the development of reliable and efficient early diagnostic platforms can potentially reduce costs, improve the quality of life, and save numerous lives. In an era with an ever-increasing dependence on technology and the availability of highly capable electronic gadgets, there lies great potential in harnessing their great computing power to solve many of these problems. In this work, three diagnostic platforms reliant on such technology have been developed to effectively detect cancer, cardiovascular disease, and speech disorders economically, accurately, and early.

Table of Contents

Dedication	iii
Acknowledgements	iv
Abstract	v
List of Illustrations	x
List of Tables.....	xvi
Chapter 1 Smart Solutions to Disease Diagnosis	1
Introduction	1
Research Overview	6
Differential Behavior of EGFR-Overexpressing Cancer Cells through Aptamer-Functionalized Micropores.....	7
Human voice, A Biometric Signature and A Biospecimen	7
Children Speech Recognition and Visual Accent Trainer	8
Smartphone based Heart Monitoring System	9
Chapter 2 Differential Behavior of EGFR-Overexpressing Cancer Cells through Aptamer-Functionalized Micropores	10
Introduction	10
Background.....	10
Materials	11
Methods	12
Micropore Fabrication using Focused Ion Beam.....	12
Mask Design	12
Silicon Wafer Fabrication	13
Batch Fabrication of Micropore	15
Aptamer Preparation	19

Preparation of anti-EGFR Aptamer Functionalized Micropores	20
Experimental Set-up	21
Selection of Target	23
Gaussian Pulse Model.....	26
Results	28
Translocation Behavior.....	30
Challenges and Solution.....	33
Single Cell Differentiation	33
New Experimental Set-up.....	33
Results for Single Cell Differentiation	35
Discussion	40
Chapter 3 Human Voice, a Biometric Signature and a Biospecimen	43
Introduction	43
Voice as a Viable Biometric.....	44
Reliability of Voice Based Authentication System.....	45
Background.....	46
Methods	49
Data Collection and Database Formation	49
Feature Extraction	51
System Verification.....	51
Results	52
Algorithm 1	52
Algorithm 2	60
Algorithm 3	63
Speaker Recognition using Split Spectrum (Algorithm 4)	66

Spectrum Splitting	66
Determination of Significant Bands.....	67
Target Identification using Split Spectrum	72
Comparison	75
Peaks of FFT Magnitude Spectrum Analysis	77
Discussion	81
Chapter 4 Children Speech Recognition and Personal Language Trainer.....	82
Introduction	82
Results and Discussion	83
Visual Accent Trainer.....	89
What is it?	89
Formant Frequency	89
Methods	90
User Guide for the Application	90
Results and Discussion	96
What Problem it Solves?	96
Comparison with Similar Existing Technologies	97
Chapter 5 Smartphone based Blood Pressure and Heart Rate Monitoring	
System	98
Introduction	98
Cardiac Arrhythmia.....	98
Hypertension	99
Background.....	102
Cardiac Cycle	102
Heart's Electrical System.....	105

Pulse Pressure	105
Methods	107
Data Collection	107
Feature Extraction	107
Results and Discussion	108
Optical Absorption	108
Pulse Rate and Pulse Pressure Measurements	111
Challenges and Solutions	113
Chapter 6	115
Summary	115
Future Scope	118
Appendix A Preparation of Ag/AgCl electrodes	119
Appendix B Isolation and Culturing of hGBM Cells	121
Appendix C Fast Fourier Transform.....	123
Appendix D Correlation Coefficient.....	125
References.....	127
Biographical Information	138

List of Illustrations

Figure 1-1 Age-adjusted Death Rates in USA [6]	2
Figure 2-1 (a) Array of square windows, (b) Confocal micrograph of a single window on a mask.....	12
Figure 2-2 (a) Confocal micrograph of a window opened after BHF etch of patterned SiO ₂ , (b) SEM micrograph of anisotropically etched silicon. Black region at the bottom of the etched groove is SiO ₂ membrane with no micropore, and (c) SEM micrograph of a drilled micropore [32].....	14
Figure 2-3 (a) Array of square window with alignment markers, (b) Array of micropores with alignment markers (inset shows the magnified version of the section)	15
Figure 2-4 (a) Square window on photoresist after development, (b) Profiles of the window measured using profilometer	16
Figure 2-5 (a) Square window on silicon dioxide after buffered oxide etch (BOE) with photoresist, (b) Profiles of the window measured using profilometer	17
Figure 2-6 (a) Micropore on photoresist (NR9-1500), (b) Micropore on silicon dioxide after buffered oxide etch.....	18
Figure 2-7 (a) SEM micrograph of the groove created by TMAH etch, (b) SEM micrograph of silicon dioxide membrane with micropore, and (c) Magnified version of the micropore	19
Figure 2-8 Protocol for aptamer attachment of silicon substrate	21
Figure 2-9 (a) Micropore chip packaging assembly (not to scale), (b) Arrangement of micropore for cell passage along with electrical biasing and measurements setup, and (c) Final assembled device [32].	22
Figure 2-10 Size distribution of hGBM cells on surface (N=108).....	24

Figure 2-11 Schematic showing interactions of hGBM cells with aptamer-functionalized and bare micropores (not to scale). (a) The tumor cells have significantly higher concentration of EGFR on the cells walls. The anti-EGFR aptamers are in hairpin form. (b) The cells interact with the aptamer. Left pane shows a normal cell that has significantly less binding with the pore-grafted aptamers than that for the tumor cell shown in the right pane. (c) The schematic of the ionic current pulses resulting from respective cell passage as in (b) [32]..... 25

Figure 2-12 A simulated pulse used to model the current pulses obtained from micropore blockage [32]..... 27

Figure 2-13 Histogram showing distribution of the amount of charge blocked by normal rat blood cells while passing through bare and aptamer functionalized micropores. The inset shows histogram of distribution of amount of charge blocked by hGBM cells while passing through bare and aptamer functionalized micropores [32]..... 28

Figure 2-14 Scatter plots of hGBM cells passage through a bare micropore (red circles) and same sized micropore that was functionalized with anti-EGFR aptamer (blue diamonds). The inset shows actual ionic current trace of a typical pulse from an aptamer-functionalized micropore of 20 μm diameter when single hGBM cell passes at a flow rate of 20 $\mu\text{l}/\text{min}$. The electrical data acquisition is done at 200 KHz [32] 30

Figure 2-15 Scatter plot of pulse widths versus current peaks of the pulses from rat blood cells passage through a bare (blue cross), and same sized micropore that is functionalized with anti-EGFR aptamer (red circle). Diameter of micropore is 20 μm , flow rate is 20 $\mu\text{l}/\text{min}$. The electrical data acquisition is done at 200 KHz. Inset shows a magnified version of the densely populated region near origin. The functionalized pore data is shown in front of bare micropore data of the main figure. The two data are dense and lie on top of each other [32] 31

Figure 2-16 (a) Experimental set-up for single cell differentiation using two micropores in series, (b) An exemplary electrical pulse	34
Figure 2-17 hGBM cells detection using two micropores in series	35
Figure 2-18 Rat blood cells detection using two micropore in series	36
Figure 2-19 Scatter plot of peak current vs pulse width of (a) hGBM cells, (b) Rat blood cells passed through two micropores in series	38
Figure 2-20 Histogram of blocked charges by (a) hGBM cells (N =173), (b) Rat blood cells (N = 100)	39
Figure 3-1 Flow chart of a speaker identification algorithm	44
Figure 3-2 Human larynx (a) Vertical cross-section, (b) Horizontal cross-section (adapted from [55, 56])	47
Figure 3-3 Average of normalized magnitude spectra of vowel sounds for (a) Subject 1, (b) Subject 2, and c) Subject 3 (n=5 for each vowel). Partial spectrum (0 – 5 KHz) is shown since most of the features lie within this region	50
Figure 3-4 (a) Time domain representation of Sound A (recording time = 2 sec, sampling frequency = 22050 Hz), (b) Normalized magnitude spectrum of that signal	51
Figure 3-5 Flow chart of Algorithm 1 (a) Identification, (b) Verification	53
Figure 3-6 Average normalized magnitude spectra (0 – 3.5 KHz) of (a) Sound “A”, (b) Sound “E”, (c) Sound “I”, (d) Sound “O”, and (e) Sound “U” of three subjects (n=5 for each average)	55
Figure 3-7 Correlation coefficient between the magnitude spectrum of test samples and their corresponding signatures for three subjects. X-axis represents test samples of a vowel pronounced by three subjects; Y-axis represents the correlation coefficient, (a) Sound “A”, (b) Sound “E”, (c) Sound “I”, (d) Sound “O”, and (e) Sound “U”	58

Figure 3-8 Correlation distance between test sample of a subject with other two subjects, (a) Subject 1, (b) Subject 2, and (c) Subject 3.....	59
Figure 3-9 Mean correlation coefficient vs pair of sounds. X-axis represents the pair of sample sounds for a particular subject used to find the mean; Y-axis represents the mean of these two correlation coefficients, (a) Subject 1, (b) Subject 2, and (c) Subject 3. Part c also explains correlation distance	60
Figure 3-10 Mean correlation distance for pair of sounds of one subject with that of others (a) Subject 1, (b) Subject 2, and (c) Subject 3.....	61
Figure 3-11 Flow chart of Algorithm 2 (a) Identification, (b) Verification.....	63
Figure 3-12 Flow chart of Algorithm 3 (identification)	65
Figure 3-13 Flow chart of Algorithm 3 (verification)	65
Figure 3-14 Spectrum splitting of a Sound (“A”) with an equal Mel interval of 250 Mel ...	68
Figure 3-15 Correlation coefficients between a sample of sound “A” of Subject 1 with three stored signatures of three subjects including Subject 1 for different frequency bands; (a) Band 1, (b) Band 2, (c) Band 3, (d) Band 4, (e) Band 5, and (f) Band 6	70
Figure 3-16 Correlation coefficients between a sample of Sound “A” of Subject 1 with three stored signatures of three subjects including Subject 1 for different frequency bands; (a) Band 7, (b) Band 8, (c) Band 9, (d) Band 10, (e) Band 11, and (f) Band 12 ...	71
Figure 3-17 Weighted average of correlation coefficients vs number of most significant bands. X-axis represents the number of most significant bands in a descending order to calculate the average and Y-axis represents the weighted average of the correlation coefficients, (a) Sound “A” of Subject 1, (b) Sound “A” of Subject 2, (c) Sound “A” of Subject 3, (d) Sound “O” of Subject 1, (e) Sound “O” of Subject 2, and (f) Sound “O” of Subject 3	73
Figure 3-18 Flow chart of the identification using split spectrum (Algorithm 4)	74

Figure 3-19 Flow chart of the verification process using split spectrum (Algorithm 4)	75
Figure 3-20 Location of peaks for (a) Sound of “A”, (b) Sound of “O”	78
Figure 3-21 Peak widths of (a) Sound of “A”, (b) Sound of “O”	79
Figure 3-22 Normalized peak amplitudes of (a) Sound of “A”, (b) Sound of “O”	80
Figure 4-1 (a) Screenshots of home screen, (b) when “Record” is pressed, a small window prompts to speak (circle). Once speaking is finished, it detects the word and shows in the textbox above, (c) Screenshots of home screen while saving a word. Once the word is detected, press “Submit” to save the word in database. The left figure shows a word (book) while saving in the database, (d) when “Last Item” is pressed, it shows what item was stored last in the textbox right below of that button	85
Figure 4-2 (a) Screenshots of the database showing all the words, (b) An index number is entered, the textbox on the right shows the indexed item, (c) Screenshots of a warning message when pressed “Clear DB”, (d) Explaining quantification of how accurately a word was pronounced	86
Figure 4-3 (a) Screenshot of home screen with highlighted “Help” button, (b) When the “Help” button is pressed, it prompts a new screen that describes the functionality of each button in the application	88
Figure 4-4 (a) Screenshot of home screen of the MATLAB graphical user interface (GUI) application, (b) Time domain signal of the sound of English vowel “A”	90
Figure 4-5 (a) Time domain signal of sound of “A”, (b) Filtered version of the same signal ($f_{CL} = 40$ Hz and $f_{CH} = 3500$ Hz), and (c) Automatically detected signal block.....	91
Figure 4-6 (a) FFT magnitude spectrum of sound of “A”, (b) FFT magnitude spectrum of the filtered version of the same signal	92
Figure 4-7 (a) Parameterization of Sound “e” trial one, (b) trial two	93
Figure 4-8 (a) Spectrogram of sound of “A”, (b) Detected signal block	95

Figure 5-1 A typical ECG signal [66].....	99
Figure 5-2 Cardiac cycle [79]	102
Figure 5-3 A simple model of light passing through a lossy medium (medium 2). Medium 1 is air.....	108
Figure 5-4 Light absorption in fingertip during cardiac cycle	110
Figure 5-5 Time variation of light intensity on a fingertip during cardiac cycle	111
Figure 5-6 Pulse rate measurement using a standard BP monitor and smart phone.....	112
Figure 5-7 Comparison of blood pressure measurement using a standard monitoring device and smartphone.....	113

List of Tables

Table 2-1 Oxidation of silicon wafer (double side polished, diameter 4")	13
Table 2-2 Statistics of electrical pulses for hGBM (blue) and rat blood (red) cells (Data collection time 10 sec, N=173, travel time 18.59 ± 2.04 ms)	40
Table 3-1 Correlation coefficient between the magnitude spectra of the test samples and signatures.....	56
Table 3-2 Relation between frequency and Mel number	66
Table 3-3 Significant bands and their weights for Sound "A" and "O"	72
Table 3-4 Memory storage, identification and verification time for 1000 users (n = 22) ...	76
Table 3-5 Comparison of correlation coefficient and correlation distance	76
Table 5-1 Classification of blood pressure for adults [68].....	100

Chapter 1

Smart Solutions to Disease Diagnosis

Introduction

We live in a world with the abundance of technologies and the usage of technology is accelerating at an ever increasing pace. Now many people from developing countries and almost everyone from developed countries possess multiple electronic gadgets. To the very least, each person carries one smart device- a cell phone. Today's cell phones, even with a moderate configuration, have more computing power than the sixth generation Intel microprocessors. All of these smart devices (cell phones, tablets, mini pcs, digital notebooks, e-readers etc.) have been underutilized in terms of their computational capacities. To maximize the capacity of these devices, applications development industries for android, windows or iOS platforms have been expanding exponentially for last few years. We can take advantage of these daily life technologies to make our lives better. What if we could turn a laptop with a data acquisition system into a disease diagnosis kit? What if a cell phone is turned into a heart monitoring system that can prevent cardiovascular diseases like hypertension or cardiac arrhythmia? What if your tablet can identify a speaker by listening his/her voice or it can monitor health of vocal cords from variations in the output of sound? What if a mobile application helps children with language impediment to learn language in a more engaging ways? If the answers of the above questions are yes, definitely it would make our lives much easier and safer. We can develop systems for the smart platforms that have already been made in bench top or desktop format. The core devices to develop such platforms are out there for a while. All we need to build are systems capable of performing particular tasks. These new systems can be thought as accessories to the existing technology that can

enable us to do more with our smart devices. The main advantages of smart systems are their easy accessibility, affordability, and portability. The impact can be huge in terms of improving the quality of lives.

Considering the mortality rate, cancer is one of the leading causes around the globe. In the United States, it is placed as the second top cause of deaths as shown in Figure 1-1. Deaths caused by some cancer types have not dropped significantly during the last six decades [1]. There are six biological capabilities that are gained during the multistep development of tumors and considered as the hallmarks of cancer. These are sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis and activating invasion, and metastasis. The actual causes of cancer are still not very well understood. However, the major hallmarks of cancer are mutations in genes, inability to repair the genes, and abnormal expression of genes [2-4]. These traits may be inherited or caused by environmental factors such as carcinogens, tobacco, radiation, etc. [5].

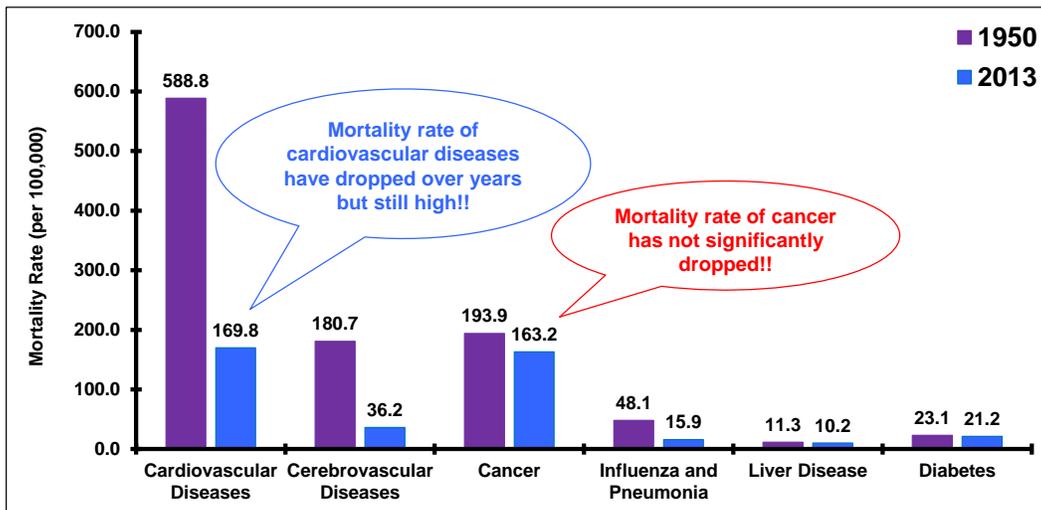


Figure 1-1 Age-adjusted Death Rates in USA [6]

Conventional treatments have very low cure rates once the tumor become metastatic. Therefore, the stage at which cancer diagnosis is done is very critical for the survival of the cancer patients. Early detection of cancer decreases the death rates of lung, breast, and colorectal cancers significantly [7]. However, this chance drops drastically for patients diagnosed at later stages. Hence, early diagnosis possibly at premalignant stage may reduce the mortality rate dramatically.

Unavailability of consistent phenotypic traits is the primary obstacle toward cancer diagnosis. Besides, current diagnosis methods are time consuming and require bulky expensive equipment. Moreover, it's almost impossible to interpret the test results without the help of an expert. A point-of-care (POC) device can solve many problems of the existing diagnostic tools. Portability, user friendliness, and minimal and noninvasive sample drawing from the patient are some of the main features of a POC device. Researchers have incorporated many features to design POC systems. Hence, current cancer research is almost exclusively multidisciplinary. However, no method has shown effectiveness and efficacy at a level required for mass scale production so far. Each method has its own limitations and most of these don't find a way outside of research lab.

Our body fights against all diseases and foreign agents through bio-affinity based interactions. This approach is proven effective; otherwise we couldn't survive in an environment which is full of parasites and infectious organisms. As a result, affinity based cancer detection schemes are very important for early cancer detection. In most cases, the actual reasons for cancer are unknown. But, it has been established that every cancer is associated with one or more mutated genes, called oncogenes. These oncogenes are responsible for abnormal protein formation (wrong sequence of amino acid, wrong folding, etc.) or unregulated protein synthesis. In any case, one or more cellular pathways are disrupted thereby facilitating uncontrolled cell proliferation. As a

result, detecting and quantifying such abnormal protein expression, known as biomarkers, is the main approach in most diagnosis. However, all these approaches suffer from high rate of false results due to their statistical natures. The level of proteins in a healthy individual may vary depending on age, heredity, and environment. The variations caused by natural factors make the boundary between a healthy and a tumor cell very ill-defined. The cells expressing biomarkers with a sufficient degree are very few in number. At very early stages, such cells with biomarkers get into the blood circulation. They become known as circulatory tumor cells (CTCs). Wandering through the circulatory system, CTCs can land into a new tissue site and start a secondary tumor. The process by which a tumor cell detaches from its origin and migrates into a new site to form secondary tumor is called metastasis. Once metastasis starts, it's virtually impossible to stop the cancer. Apart from the exceedingly low number of CTCs (a few in a million) in peripheral blood, the situation is further exacerbated by the fact that the number of blood cells in a healthy individual varies over a wide range. As a result, making a device that can detect and quantify CTCs with great sensitivity and selectivity is complex and daunting.

Most of the current approaches for cancer cell detection and isolation rely on target tagging, mechanical behavior monitoring, and affinity based interactions between a probe and a biomarker. Factors like mechanical forces [8, 9], dielectrophoresis [10], optical interactions [11], immunohistochemistry [12], magnetic sorting [13], flow cytometry [14], etc. have been reported to identify and isolate cancer cells from peripheral blood. Affinity based interactions provide higher efficiency and greater specificity compared to mechanical and electrical sorting techniques [15]. Affinity based molecules such as antibodies use key-lock mechanism to arrest the target molecules. Our body neutralizes foreign agents using the same mechanism. The use of antibodies is, however, subject to

high levels of off-target cross-reactivity [12, 16, 17]. Another challenge of using antibodies for detection is that it is very difficult to retain specific sample conditions that are required for antibodies to be completely functional. Many technical challenges related to reproducibility are also encountered when cross-linking antibodies onto the surfaces of miniaturized devices due to non-homogeneity of conjugation and denaturation on the surface. Recently, a special type of single stranded DNA or RNA sequence called aptamer has been shown to have similar affinities and specificities like antibodies [18]. The aptamers are very stable at a variety of salt and ionic conditions and can be reversibly denatured [19, 20]. In recent works, aptamers have been used to activate cell signaling pathways and also to label cells [21-24]. Devices incorporating aptamers to sort, isolate, and detect tumor cells have also been reported [18, 25-28]. The main advantages of aptamers over antibodies lie in the ease of chemical synthesis, site-specific labeling, and, therefore, site-specific immobilization.

All the efforts made toward early cancer detection can be classified into two broad categories:

- 1) Surface chemistry based detection. Selectivity of this approach is very good, but sensitivity is poor.
- 2) Micropore based detection. It offers great sensitivity (single cell level), but selectivity is based on physical size only

There are many diseases that can be prevented if proper monitoring devices are employed at right time. Heart diseases, also known as cardiovascular (CV) diseases refer to any disease that involves the heart, the blood vessels (arteries, capillaries, and veins) or both. In last six decades, the death toll due to CV diseases has reduced from 588.8 to 169.8 per hundred thousand of the population as shown in Figure 1-1 [6]. But it is still the leading cause of fatalities in US since 1950. The huge reduction in fatalities is attributed

to the development of many screening devices, better medication, minimally invasive surgical procedures, and awareness. There are many cardiovascular diseases in human such as myocardial infarction (heart attack), sudden cardiac death (SCD), angina, coronary artery disease (atherosclerosis), congestive heart failure, cardiomyopathy, cyanotic heart diseases, hypertensive heart disease (high blood pressure), inflammatory heart disease, heart failure, arrhythmia, endocarditis etc. It has been widely accepted that most of the cardiovascular diseases except congenital ones can be prevented through better life style (dietary habits, exercise, regular monitoring, etc.) [29]. Since heart is responsible for pumping blood to the whole body and it does so in a periodic manner, most of the cardiovascular diseases, especially hypertension and arrhythmia affect the blood pressure or the heart rate or both. A method has been developed on a smart platform for monitoring heart rate (HR) and blood pressure (BP) that prevents these CV complications from escalating into further detrimental states.

Research Overview

The primary focus of this dissertation is to develop inexpensive diagnostic platforms for diseases, especially cancer and cardiovascular diseases detection. Chapter 2 studies the behavior of epidermal growth factor receptor (EGFR)-overexpressing cancer cells through aptamer-functionalized micropores. This approach exploits the aptamer-protein interactions that facilitate early cancer detection. Chapter 3 discusses a simple text-dependent speaker recognition method using English vowels. The method has the potential to provide a noninvasive way to monitor our vocal cords using speech analysis. Speech signal is recorded as sample from the patient on a regular interval and analyzed to find if the patient is building any polyp or nodule on his/her vocal cords. Chapter 4 concentrates on developing a speech recognition system for children with

speech impediment and a visual accent trainer for a foreign language learner. Speech and language therapists may use the application to evaluate the speech production of each patient quantitatively. Visual accent trainer helps foreign language learners to get rid of the accents by providing visual feedback of their speech production. Chapter 5 focuses on a smartphone based health monitoring system that can measure heart rate and blood pressure in a convenient way. Chapter 6 is appendix that includes the details of several processes used in the research.

Differential Behavior of EGFR-Overexpressing Cancer Cells through Aptamer-Functionalized Micropores

Cancer cell overexpresses several biomarkers for different types of cancer. Some well-known biomarkers are EGFR for brain and lung cancer, HER2/neu for breast cancer, gastric cancer, and esophageal cancer; CA-125 for ovarian cancer, CA 15-3 for breast cancer, CEA for colon cancer, and PSA for prostate cancer etc. EGFR was targeted as a biomarker for human glioblastoma (hGBM, brain tumor). An RNA molecule (anti-EGFR aptamer) that can selectively bind to EGFR was used to functionalize a micropore. Rat blood and hGBM cells were passed through functionalized and non-functionalized micropores while measuring the ionic current through the pores. The result showed distinct profiles for the functionalized micropore. The translocation behavior analysis of cells through aptamer functionalized micropore is a new modality to detect cancer at early stage.

Human voice, A Biometric Signature and A Biospecimen

Speech is the output of a well-coordinated muscular activity. It is the prime tool for human communication. Speech carries two types of information- message and information about speaker. The information about the speaker contains the identity of the speaker and the conditions of his/her sound-producing apparatus such as vocal chords,

tongue, laryngeal tube, etc. As a result, speaker identity as well as the health condition of sound-producing apparatus can be extracted by analyzing speech. Human voice has potential to be a viable biometric like fingerprints. An authentication system has been developed to identify speakers using their voices as signatures. Sounds of English vowels from subjects were taken as inputs and processed to extract a feature from each sample. These features were stored in a database. Magnitude spectrum of recorded sound was used as the feature in this study. The authentication was done by computing the percentage match of the sample's feature with that of the stored in the database. When the matching percentage was higher than a threshold, it was considered as a hit. The system also has the potential to monitor health of sound producing apparatus non-invasively.

Children Speech Recognition and Visual Accent Trainer

Language is one of the sophisticated invention of human as a species. This is the best form of communication medium by far. We learn it at very early stage of our lives. Not all children learn their native languages at a same rate. Some of them take longer than others because learning to speak a language takes a large amount of information processing and coordinated muscular activities. As a result, many children face language learning difficulties and there's no reliable tool in the market that can help them. Currently, a speech and language therapist listens to each word produced by a kid and evaluates its performance and then decides what therapy the kid might need. Therapist needs to record the sound and the accuracy of the vocalization manually to a computer. An android application has been developed to alleviate this problem. Parents may use the developed application to help their kids or it may complement the services of speech and language therapists or other existing applications. It transcribes each word produced by a child undergoing speech therapy and tracks them. Words produced by each child is

analyzed to compute the accuracy of the pronunciation by providing feedback. The application not only tells a patient that the sound production was not accurate but also provides a quantitative feedback about how accurate it was. Since each sound is produced from a specific location of the sound-producing apparatus, the application offers a non-invasive way to investigate the source of the speech disorder. The word delivery by each child can be recorded and quantified to measure progress automatically.

Visual accent trainer was developed on MATLAB GUIDE to help new language learner and hearing impaired people to produce sounds more accurately without any accents. It provides a visual quantitative feedback about the sound produced by comparing with a pool of native speakers. The application is more like a target practicing game where you can see your target and practice until the target is reached.

Smartphone based Heart Monitoring System

Heart diseases, also known as cardiovascular (CV) diseases are the leading cause of death in USA. Due to the development of many screening devices, the fatalities have dropped significantly during the last six decades. The huge reduction in fatalities is attributed to the development of many screening devices, better medication, minimally invasive surgical procedures, and awareness.

A heart monitoring system capable of measuring heart rate and blood pressure using a smartphone has been developed. With this convenient system, people can have an easy way to monitor their hearts without going to a clinic. Some of the pathological conditions of heart may not be felt by the patients at early stage. But with a convenient and affordable monitoring system, gradual changes of heart rate and BP can be detected before it's too late. It can be used as a primary diagnostic tool that can help the patient to decide whether s/he needs to see a doctor or not.

Chapter 2

Differential Behavior of EGFR-Overexpressing Cancer Cells through Aptamer- Functionalized Micropores

Introduction

An aptamer functionalized micropore system was developed to study behaviors of tumor cells. Micropore based detection system has been well known as cell or particle counter due to its sensitivity [26, 30]. Its detection limit can go as low as single cell level. On the other hand, aptamer functionalized surfaces have been reported to capture, isolate or target tumor cells due to the high affinity based interactions between aptamer and the tumor cells [27, 28, 31]. Our method brings the selectivity of the aptamer functionalized surface together with the sensitivity of micropore based detection system. As a result, it has both the sensitivity and selectivity that are highly desired for an effective cancer detection system. It has been reported that tumor cells (hGBM) behave differently from normal cells when these were pumped through anti-EGFR aptamer functionalized micropore [32]. Their translocation behaviors were distinct from the normal counterpart. The differentiation was attributed to the selective interaction between the anti-EGFR aptamer and the receptors (EGFR) on the cell membranes. This result indicated that aptamer functionalized micropore offered a new modality of cancer detection without the requirement of cell tagging.

Background

When a healthy cell turns into a tumor cell, some changes occur in the level of expression and structure of few membrane proteins. Mutations in genetic codes are primarily responsible for these changes [2-4]. Eventually, gene mutation leads to over or

underexpression or wrong folding of proteins. Wrong sequenced or folded protein as well as the expression level of a protein can be used as potential biomarkers for specific type of cancers. These are the precursors of an imminent cancerous condition. These can foretell that some cells will become malignant soon. Many biomarkers have been identified over the years. For example, HER-2 and EGFR are biomarkers for breast cancers [33, 34]; EGFR, KRAS, BRAF for lung cancers [35-37]; EGFR, HSP27, Annexin for bladder cancer, etc. [38-40]. Among these biomarkers, EGFR is the one that is overexpressed in most type of cancers. Anti-EGFR antibody and aptamer both can be used to capture or bind to tumor cells with overexpressed EGFR. However, aptamer based approach has some inherent advantages over antibody based method. Aptamers are more stable in a wide range of ionic and salt conditions. The selectivity of aptamer is better than antibody and it is much more tolerant to variations in the physiological environment. An anti-EGFR RNA aptamer was used in the study due to the benefits it offers over antibody.

Materials

Silicon wafers were bought from University Wafer Inc. Aptamer was provided by collaborator from UT-Austin. Silicon etchant, tetramethylammonium hydroxide (TMAH) was purchased from VWR International. The linker amino modified DNA oligo was purchased from Integrated DNA Technologies (IDT). RNase free was bought from Argos Technologies Inc. Silver (Ag) wires with 0.375 mm diameter were bought from Warner Instrument. Chloriding solution (Clorox) was bought from Walmart to coat Ag wires. Appendix A provides the details of Ag/AgCl electrode preparation. Rat's blood was obtained from 6 month old rat (*Sprague Dawley*) which was purchased from Charles River. It was collected from the rat's tail by restraining the rat. The blood was then stored

in K₂-Ethylenediaminetetraacetic acid (EDTA) tube to prevent coagulation. Blood samples were stored in refrigerator at 4 °C. The tumor cells (hGBM) samples were collected from consenting patient at the University of Texas Southwestern Medical Center (Dallas, TX). The isolation and culturing protocols are delineated in Appendix A [18]. The remaining chemicals used in the experiments were purchased from Sigma Aldrich.

Methods

Micropore Fabrication using Focused Ion Beam

Mask Design

Every fabrication process starts with designing a mask. A single layer mask (5" x 5") was designed using AutoCAD and printed on glass substrate for micropore fabrication as shown in Figure 2-1. It has an 8 x 8 array of square windows. These windows were used to create openings in silicon dioxide membranes. Once the mask was designed and fabricated, then silicon wafer fabrication was started.

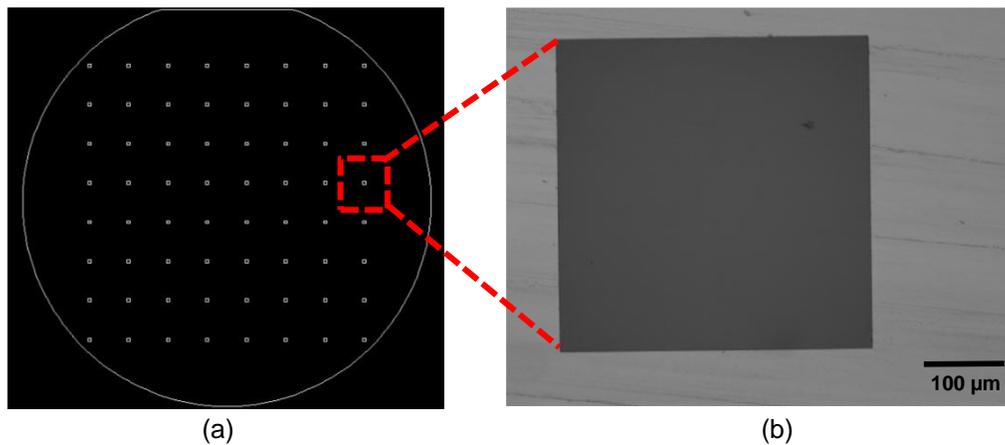


Figure 2-1 (a) Array of square windows, (b) Confocal micrograph of a single window on a mask

Silicon Wafer Fabrication

At the first step of wafer fabrication, 0.9 μm thick silicon dioxide (SiO_2) layers were grown on both sides of a double-side polished (100) p-type silicon wafer. The oxide thickness of a batch of silicon wafers was measured using Reflectometer (Ocean Optics NC-UV-VIS Reflectometer) and listed in Table 2-1.

Table 2-1 Oxidation of silicon wafer (double side polished, diameter 4")

	Wafer 1		Wafer 2		Wafer 3		Wafer 4	
	Front Side (nm)	Back Side (nm)						
1	860.00	860.40	865.50	867.30	867.90	868.50	867.90	867.60
2	868.30	867.30	871.10	873.20	873.60	871.10	872.50	872.00
3	869.80	870.90	874.50	873.60	874.90	875.30	873.20	873.80
4	867.80	869.30	871.80	871.40	873.30	872.60	872.90	872.90
5	864.90	864.30	869.60	870.20	870.30	870.70	871.70	870.20
Avg	866.16	866.44	870.50	871.14	872.00	871.64	871.64	871.30
STD	3.47	3.74	2.96	2.28	2.54	2.25	1.94	2.20

An array of square windows was opened by photolithography to start etching on the back side of the wafer. Pattern was transferred from the mask onto wafer using standard photolithography. Once the photoresist was patterned, buffered hydrofluoric acid (BHF) was used to transfer the same pattern from the resist onto SiO_2 layer underneath as shown in Figure 2-2a. After oxide patterning, 25% TMAH solution, an anisotropic wet etchant, was used to etch silicon from the back side through these oxide windows at 90 °C. A magnetic stirrer was used at 200 rpm to ensure uniform etch. The silicon etch was self-limiting due to the presence of oxide layer on the other side of the wafer since TMAH has very high selectivity for silicon over SiO_2 . The membranes with 70 μm X 70 μm

dimension were created at the end of silicon etch. The membranes (black region in Figure 2-2b) were inspected under scanning electron microscope (SEM) before further processing. Micron sized pores were drilled on those oxide membranes using a focused ion beam (FIB, ZEISS 1540XB) machine. Diameter of a micropore (Figure 2-2c) was controlled by varying acceleration voltages, milling currents, and drilling times [41]. In order to drill a micropore with 20 μm diameter in a 0.9 μm thick SiO_2 membrane, 30 kV of acceleration voltage and 1 nA of drilling current were applied for 5 min. All the steps involved in the fabrication till FIB drilling are batch processes i.e. high throughput. But FIB can drill one pore at a time. Since any process speed is determined by its slowest step, this method of micropore fabrication is a serial process which limits the throughput of the entire fabrication. In order to overcome this limitation, a fabrication method was developed which is a batch process that doesn't require any FIB drilling.

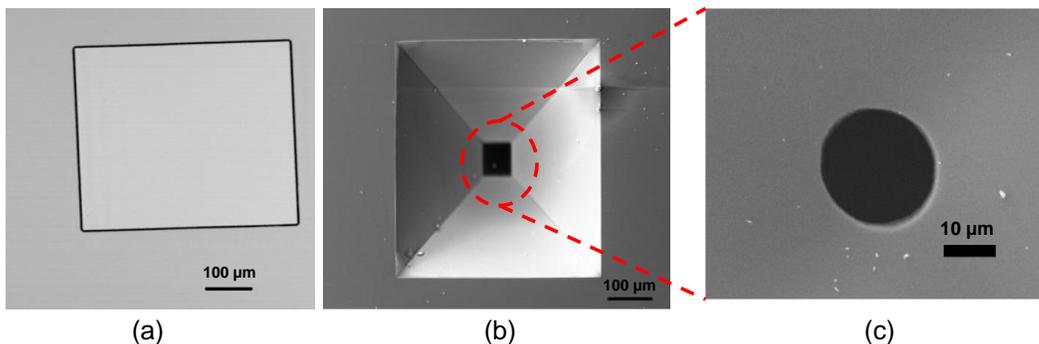


Figure 2-2 (a) Confocal micrograph of a window opened after BHF etch of patterned SiO_2 , (b) SEM micrograph of anisotropically etched silicon. Black region at the bottom of the etched groove is SiO_2 membrane with no micropore, and (c) SEM micrograph of a drilled micropore [32]

Batch Fabrication of Micropore

This method of micropore fabrication requires two masks- one for opening the square etch window which is similar to the mask used for previous method and the other for creating array of micropores on the center of oxide membranes left after the TMAH etch. Two masks are shown in Figure 2-3. Similar to the first approach, 0.9 μm thick silicon dioxide layers were grown on both sides of a double-side polished (100) p-type silicon wafer. Photolithography was carried out to create an array of square windows on the back side of the wafer.

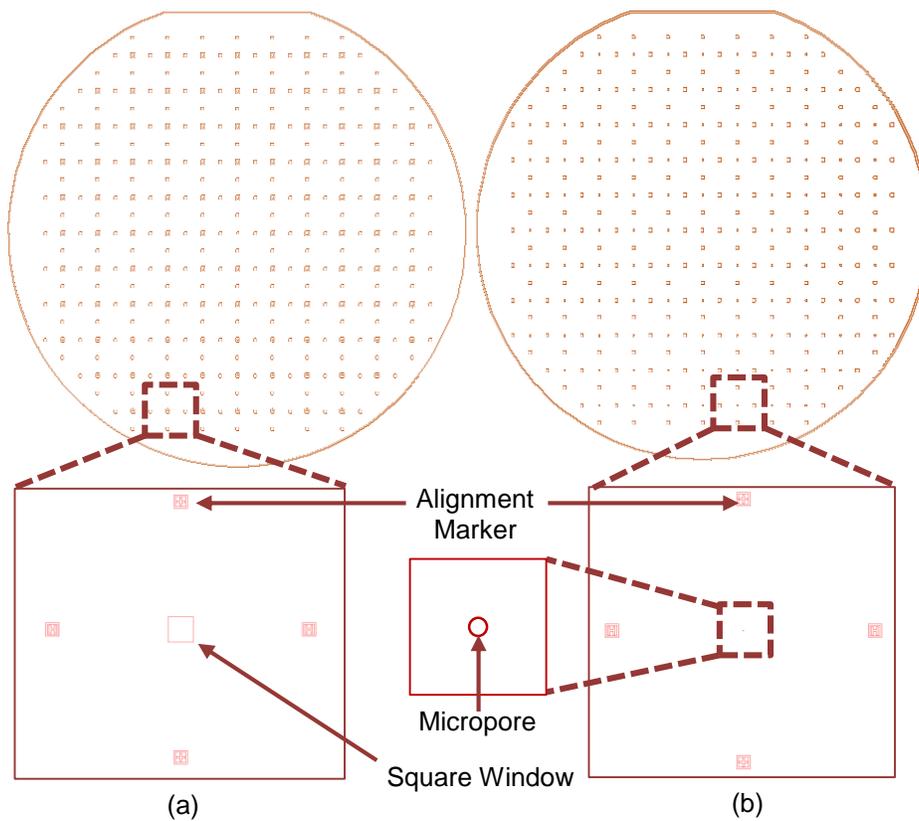


Figure 2-3 (a) Array of square window with alignment markers, (b) Array of micropores with alignment markers (inset shows the magnified version of the section)

Once the photoresist was exposed and developed, the patterns were successfully transferred from mask to the wafer. Figure 2-4a shows confocal micrograph of such window on photoresist (S1813) and Figure 2-4b shows the profiles measured in an profilometer (KLA-Tencor Alpha-Step IQ Profilometer). The pink region is the silicon dioxide and the greenish region surrounding the oxide is photoresist. The pattern was clean and edges were sharp that is the indication of successful pattern transfer.

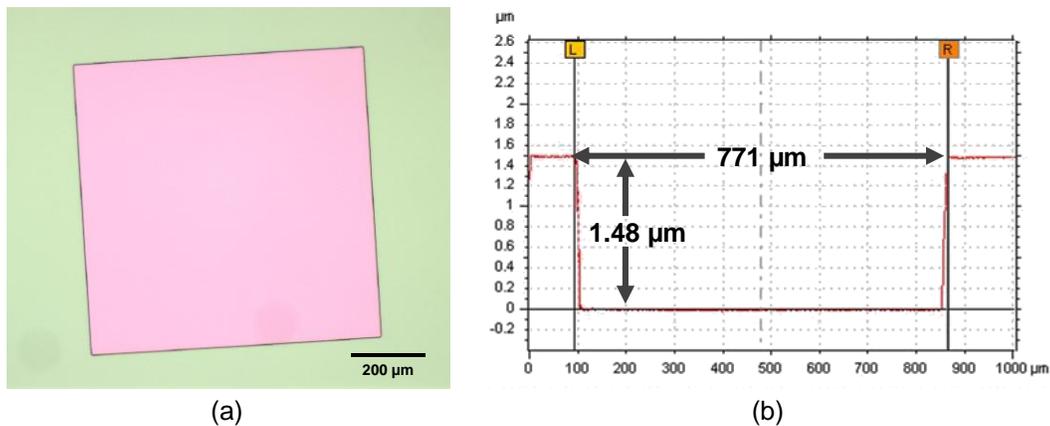


Figure 2-4 (a) Square window on photoresist after development, (b) Profiles of the window measured using profilometer

After the pattern transfer onto photoresist, oxide etch was done using buffered hydrofluoric acid (BHF). This was to open window on the underneath oxide layer. Photoresist was used as etch mask to confine the etch process only within the windows. Figure 2-5a shows the oxide window after the BHF etch. The white color is due to the complete depletion of oxide which was colorful and reflective. The photoresist was spun at 3000 rpm and the thickness it produced was 1.46 μm . With the photoresist and a 0.9 μm oxide made a thickness of around 2.36 μm which was very close to the thickness found as shown in Figure 2-5b. The window size was widened up a little (from 771 μm to

784 μm) in this step. The window size was measured with both confocal microscope and profilometer and the readings matched well.

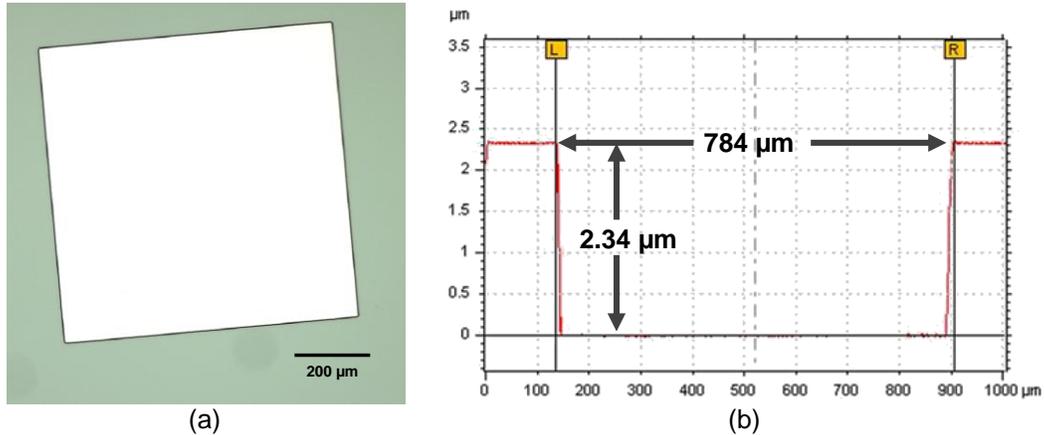


Figure 2-5 (a) Square window on silicon dioxide after buffered oxide etch (BOE) with photoresist, (b) Profiles of the window measured using profilometer

The next step was to create array of micropores aligning with the array of windows already created on the back side. The micropores were successfully aligned with the square windows using the backside aligner (OAI Model 806 manual Front/Backside Contact Mask Aligner). Once the alignment was done, the wafer was exposed and developed. A negative mask with a negative photoresist (NR9-1500py) was chosen for good visibility of the feature. Figure 2-6a shows the micropore on photoresist after development and Figure 2-6b shows the same micropore after BHF etch. The diameter of the micropore was 18 μm . The pink region on the left of Figure 2-6 shows the color of oxide and grayish region is the negative photoresist NR1500. After the oxide strip, the exposed region looks white as the shiny oxide layer is gone from that part.

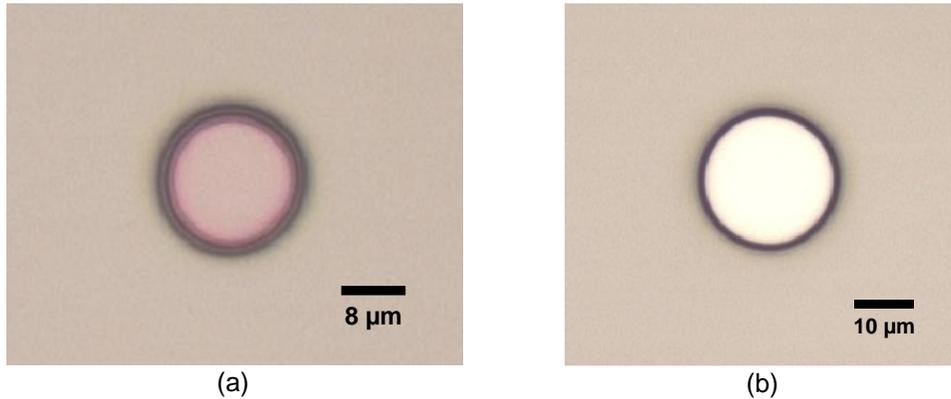


Figure 2-6 (a) Micropore on photoresist (NR9-1500), (b) Micropore on silicon dioxide after buffered oxide etch

Once the array of micropores was transferred on the oxide layer aligned with the oxide windows on the back side, the wafer had to go through a silicon etch process. To etch silicon, a solution of 25% of tetramethylammonium hydroxide (TMAH) and deionized (DI) water in a ratio of 24:11 (2400 ml TMAH with 1100 ml DI water) was used. Right before putting the wafer into TMAH solution, it was dipped into hydrofluoric acid (HF) for few seconds to remove the native oxide layer. The solution was kept at a constant temperature of 90 °C with a stirring rate of 200 rpm. Stirrer kept the solution temperature reasonably uniform throughout the entire beaker. Magnetic stirring rod was used to decouple the system. At this concentration and temperature, the etch rate of silicon along (100) plane was 0.84 μm/min. TMAH is an anisotropic etchant. It etches (111) plane 37 times slower than the (100) plane. As a result, the etch structure was like an inverted pyramid as shown in Figure 2-7a. To etch the whole thickness of a 4" silicon wafer, it usually takes approximately 10 hrs. The stirring rate was kept 200 rpm for first 7 hrs and then reduced to 100 rpm in order to reduce the damage on the oxide membrane due to mechanical agitation. The wafer was inspected several times during the etch process. Once TMAH etch was done, the wafer was inspected under optical and electron

microscopes. Figure 2-7b shows the membrane with micropore and Figure 2-7c shows the magnified image of a micropore. The final diameter of micropore was 19 μm .

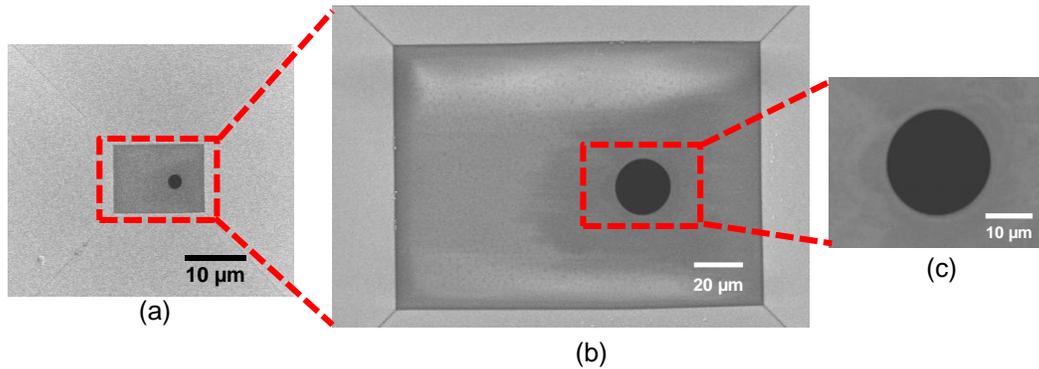


Figure 2-7 (a) SEM micrograph of the groove created by TMAH etch, (b) SEM micrograph of silicon dioxide membrane with micropore, and (c) Magnified version of the micropore

Aptamer Preparation

Protocol for aptamer preparation was adopted from an early report [42]. The sequence of the extended anti-EGFR aptamer was 5'-GGCGCUCCGACCUUAGUCUCUGUGCCGCUAUAAUGCACGGAUUUAAUCGCCGUAGAAAAGCAUGUCAAGCCGGAACCGUGUAGCACAGCAGAGAAUUAAAUGCCCGCCAUGACCAG-3'; sequence of a mutant aptamer was 5'-GGCGCUCCGACCUUAGUCUCUGUCCCCACAUCAUGCACAAGGACAAUUCUGUGCAUCCAAGGAGGAGUUCUCGGAACCGUGUAGCACAGCAGAGAAUUAAAUGCCCGCCAUGACCAG-3' (extension sequence is underlined). An amino-modified capture oligonucleotides, 5'-AmMC6-TTTTTTTTTTTTTTTTTTTTCTGGTCATGGCGGGCATTTAATTC-3' (active capture sequence is underlined) was incorporated on the surface to capture the aptamer.

Preparation of anti-EGFR Aptamer Functionalized Micropores

The protocol for aptamer attachment on the SiO₂ micropore surfaces was adapted from earlier works [27, 28, 43]. Each diced silicon chip with micropore was 8 x 8 mm² in size. Those were cleaned in piranha solution (H₂O₂/H₂SO₄, 1:3 v/v) for 10 minutes. After rinsing with deionized (DI) water and drying in nitrogen flow, the substrates were immersed in 2% v/v 3-aminopropyl tri-ethoxysilane (APTES) in ethanol for 30 minutes at room temperature. That step was to create amine groups on surface. The APTES solution was removed and the substrates were sequentially washed with ethanol and DI water. Then 2 mg of p-phenylene diisothiocyanate (PDITC) was mixed in a solution of dimethyl sulphoxide (DMSO) and pyridine (9:1 v/v) to prepare PDITC solution. The APTES treated substrates were incubated in PDITC solution at 55 °C for 5 hours. After 5 hours, the PDITC solution was removed and the substrates were washed three times each with isopropyl alcohol (IPA) and diethylpyrocarbonate (DEPC) treated DI water. Then 10 µM capture DNA solution was added to DMSO to make 1:1 (v/v) DNA: DMSO solution. The substrates were immersed in DNA solution and incubated at 45 °C overnight. The chips were then soaked in ethanol followed by sequential washing with IPA and DEPC treated DI water. Nitrogen was used to dry the chips. RNA aptamer (1 µM) was mixed in hybridization buffer at a ratio of 1:5 (v/v). Then a hybridization chamber was cleaned with RNase-free DI water thoroughly and a drop of aptamer solution was placed on the chip. The chamber was immersed in a water bath (BW-10G, Laboratory Companion) for one hour at a temperature of 37 °C. The whole process is summarized in Figure 2-8.

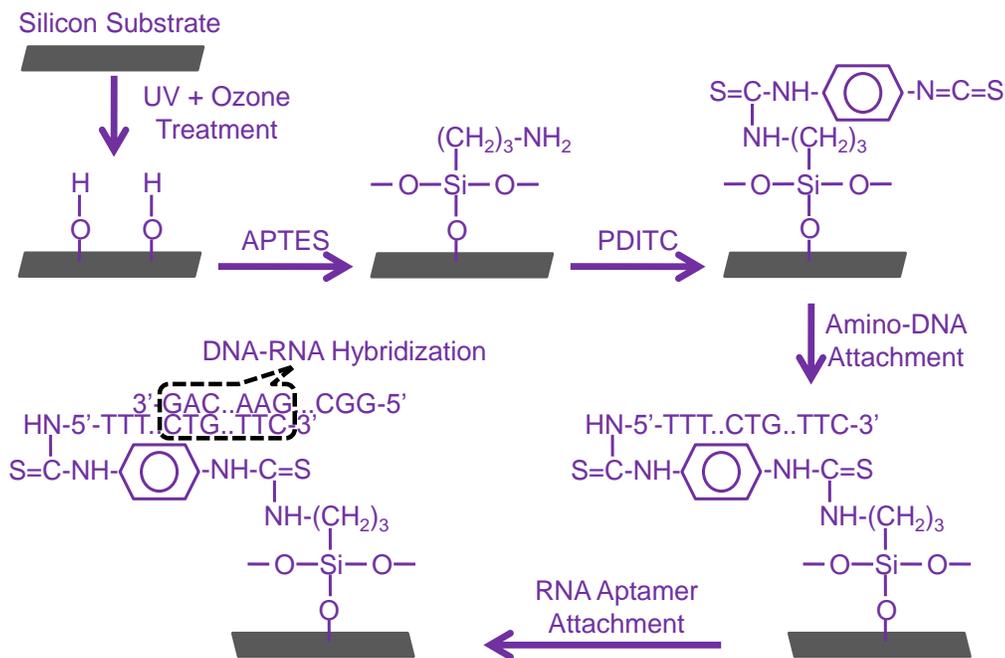


Figure 2-8 Protocol for aptamer attachment of silicon substrate

The chips were then washed again with IPA and DEPC water three times each. The 6-amino-1-hexanol solution was used to deactivate amine groups that were not bound to the nucleic acids. This solution was made by dissolving 1.1719 gm of 6-amino-1-hexanol in 5.2253 ml N,N-diisopropylethylamine and 200 ml of N,N-dimethylformamide. Chips were incubated in the deactivation solution at 45 °C for 5 hours. The chips were then washed with ethanol and DEPC treated DI water. Without drying, the chips were put in 1X phosphate-buffered saline (PBS) solution. PBS solution was heated to 45 °C for 10 minutes and then slowly cooled down to room temperature.

Experimental Set-up

The micropore chip was sandwiched between two polydimethylsiloxane (PDMS) gaskets. These gaskets were used to prevent leakage of solution. The chip and gaskets

were then held together using two Teflon blocks that had holes at their centers as shown in Figure 2-9a. Inlet and outlet tubes were connected to the Teflon blocks. Two Ag/AgCl electrodes were connected to the inlet and outlet respectively as shown in Figure 2-9b. A syringe pump (Harvard Apparatus) was used to maintain a constant flow rate of cell solution as shown in Figure 2-9c.

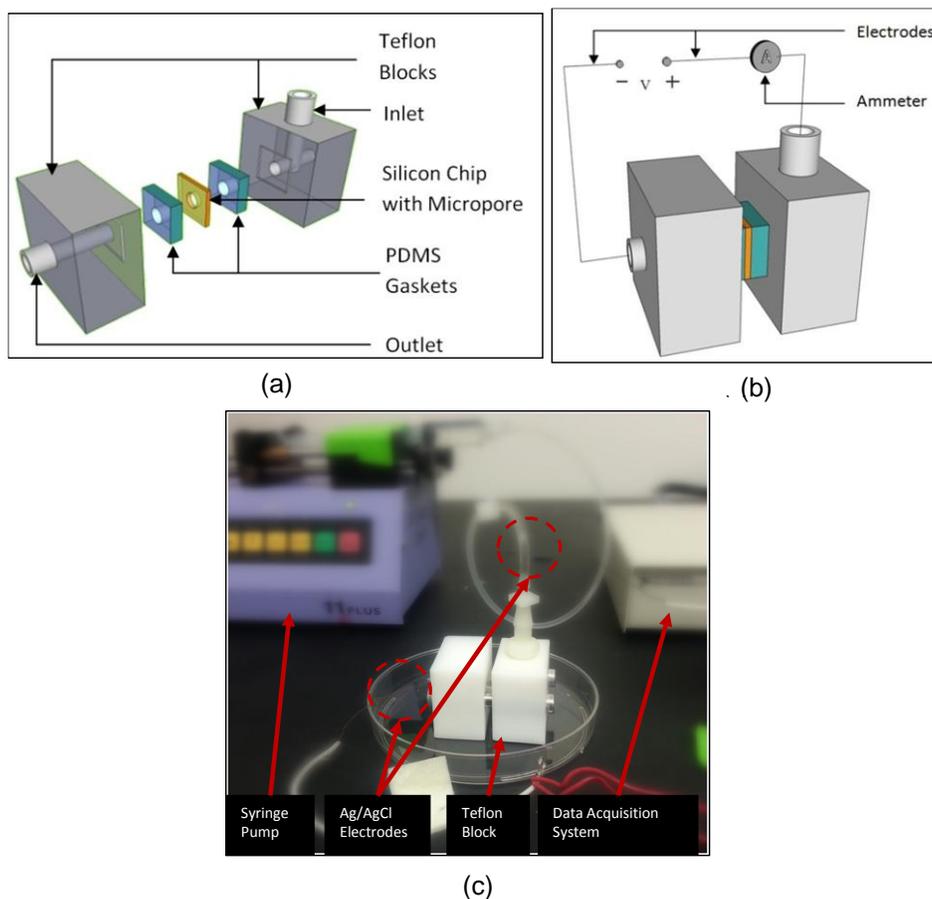


Figure 2-9 (a) Micropore chip packaging assembly (not to scale), (b) Arrangement of micropore for cell passage along with electrical biasing and measurements setup, and (c) Final assembled device [32].

The hGBM cells (5000 cells/ml) were suspended in 1X PBS and pumped through the system while measuring the current across the micropore using a data acquisition

system (National Instrument data acquisition system, NI SCB-68, NI PXI-1033, NI PXI-4071). The rat blood, collected in tubes with K₂-EDTA and stored at 4 °C, was taken out from storage tubes right before the experiments and thawed to room temperature. Then 1X PBS solution was mixed with rat blood to dilute it one thousand times. Once the blood was homogenously mixed with PBS, the solution was also pumped through micropores while measuring ionic current through them. A pair of Ag/AgCl electrodes was used to measure the ionic current. First, only PBS was pumped through the micropore to establish the baseline. The PBS gave a constant ionic current through the micropores. When the cell solution was pumped through the micropore, each cell blocked the pore temporarily when it passed through. Current dips were seen in the ionic current traces at that time. Every downward peak corresponded to a single cell passage. The hGBM cells were passed through a bare micropore and an aptamer functionalized micropore to observe the differences.

Selection of Target

For the experiment, human glioblastoma cell (hGBM) was chosen as a tumor cell line. Among many other tumor biomarkers, epidermal growth factor receptor (EGFR) was the target biomarker for the study. The level of EGFR expression in normal cell varies from 40 K to 100 K per cell [44]. It has been reported that such overexpression of EGFR especially occurs in case of lung cancer and brain cancer (human glioblastoma) [45]. Mutation in EGFRvIII acts as a continuous signaling for cell division, and uncontrolled cell division is known as cancer. An absence of a part of amino acid sequence of exon 2-7 (amino acid 6-273) causes the EGFRvIII mutation [43]. Both types of EGFRs are present on a tumor cell but the expression level of the wild type EGFR is much lower than that of mutant EGFR. It has been reported that anti-EGFR RNA aptamer binds specifically with

both mouse derived wild type EGFR and mutant EGFRvIII. Wild type and mutant EGFRs both have extracellular binding domain III. Anti-EGFR aptamer can bind to this site. As a result, aptamer has affinity for both types of EGFRs.

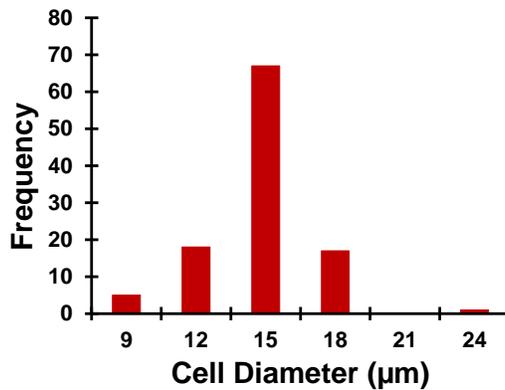


Figure 2-10 Size distribution of hGBM cells on surface (N=108)

The specificity of the aptamer to a particular site reduces any non-specific adsorption. Besides, tumor cells are more negatively charged compared to normal cell due to the higher level of sialylation on their surface [46]. The size of hGBM cells were measured in optical microscope. The size distribution is shown in Figure 2-10. The diameter of hGBM cells were $13.57 \pm 2.26 \mu\text{m}$.

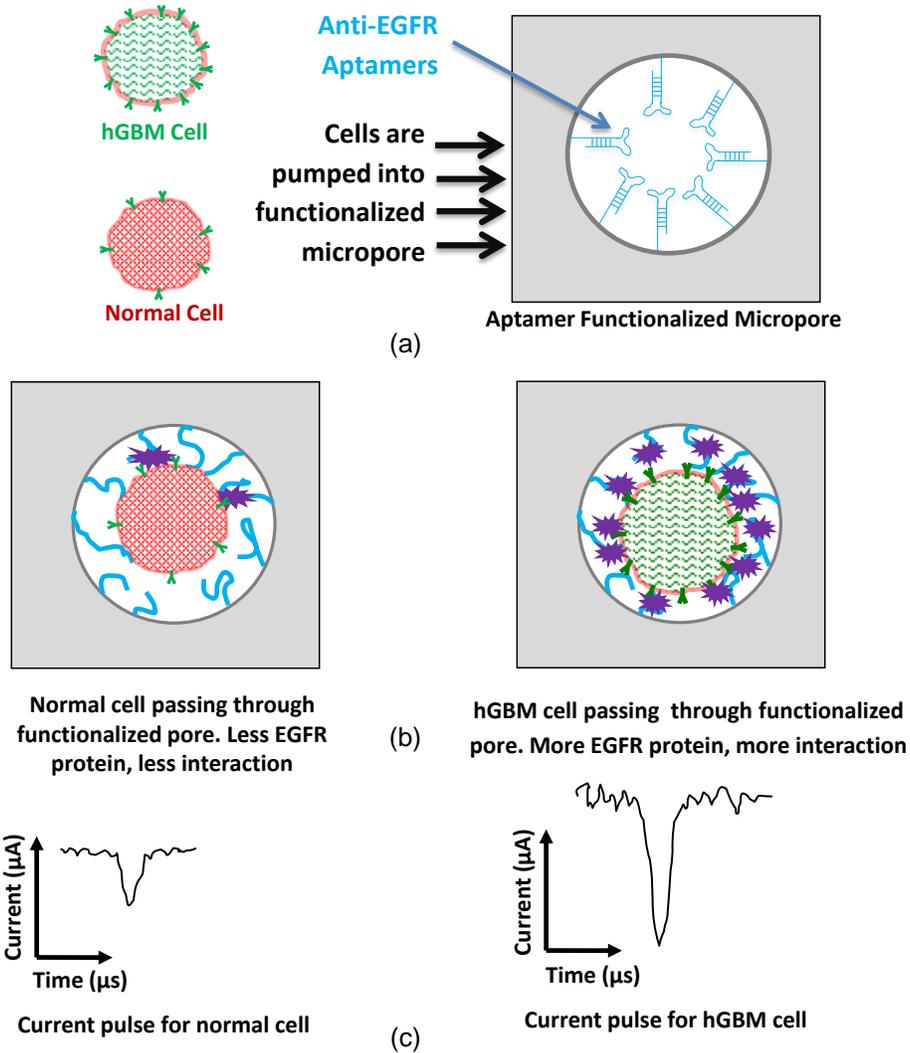


Figure 2-11 Schematic showing interactions of hGBM cells with aptamer-functionalized and bare micropores (not to scale). (a) The tumor cells have significantly higher concentration of EGFR on the cells walls. The anti-EGFR aptamers are in hairpin form. (b) The cells interact with the aptamer. Left pane shows a normal cell that has significantly less binding with the pore-grafted aptamers than that for the tumor cell shown in the right pane. (c) The schematic of the ionic current pulses resulting from respective cell passage as in (b) [32]

This extra negative charge also helped to minimize non-specific binding. The selectivity and sensitivity of aptamers were further improved by an amino modified DNA as a capturing agent. The capture DNA did covalently bind to the aptamer. It kept the aptamer in a distance from the surface which essentially reduced any steric hindrance. The overview of the detection system is shown in Figure 2-11.

Gaussian Pulse Model

The data of pulses obtained from micropores had a number of parameters that were used to understand the biophysical interactions occurring within the confinement of the micropore. When a cell passed through the micropore, it replaced or blocked the ionic species or charges from the micropore volume. That charge was then calculated. The following assumptions were made to calculate the amount of charge physically blocked/replaced by a cell while passing through the micropore.

- i) All current pulses were inverted Gaussian in shape.
- ii) Translocation time of the cell or current pulse width was assumed to be the full width at tenth of maxima (FWTM).
- iii) The micropore was perfectly circular.
- iv) Ionic concentration of the PBS was homogenous and uniform. PBS consisted of 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄ and 1.47 mM KH₂PO₄.
- v) Cells were perfectly spherical.

The equation of an inverted Gaussian current pulse is (Figure 2-12):

$$I(t) = I_{peak} \exp\left(-\frac{t^2}{2c^2}\right) \text{ for } -P_w/2 < t < P_w/2 \quad (\text{Equation 2-1})$$

Where $I(t)$ is current at time t within the pulse, I_{peak} is the peak current, and c is a parameter that controls the width of the pulse (P_w).

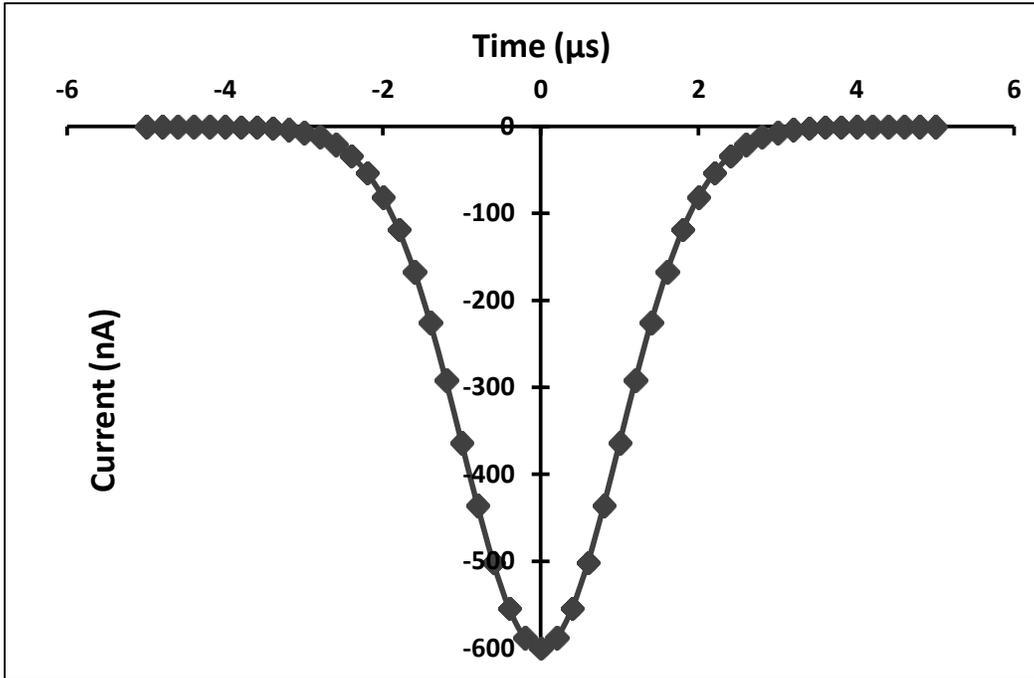


Figure 2-12 A simulated pulse used to model the current pulses obtained from micropore blockage [32]

The area under a Gaussian pulse as shown in Figure 2-12 was calculated. It was the amount of charge replaced and blocked by each cell. Now the area under the pulse, or in other words, the charge blocked, was calculated by integrating the Gaussian function with respect to time.

$$Q_{blocked} = \int_{-\infty}^{\infty} I_{peak} \exp\left(-\frac{t^2}{2c^2}\right) dt = cI_{peak}\sqrt{2\pi} \quad (\text{Equation 2-2})$$

$$\text{and } P_w = FWTM = c\sqrt{8\ln(10)} \quad (\text{Equation 2-3})$$

$$\text{From Equation 2-2 and Equation 2-3, } Q_{blocked} \approx 0.584I_{peak}P_w \quad (\text{Equation 2-4})$$

Results

Since every cell corresponded to a current pulse i.e. every cell was essentially blocking charge flow while passing the micropore. This charge was accurately depicted by the current peak (I_{peak}) and pulse width (P_w) measured experimentally.

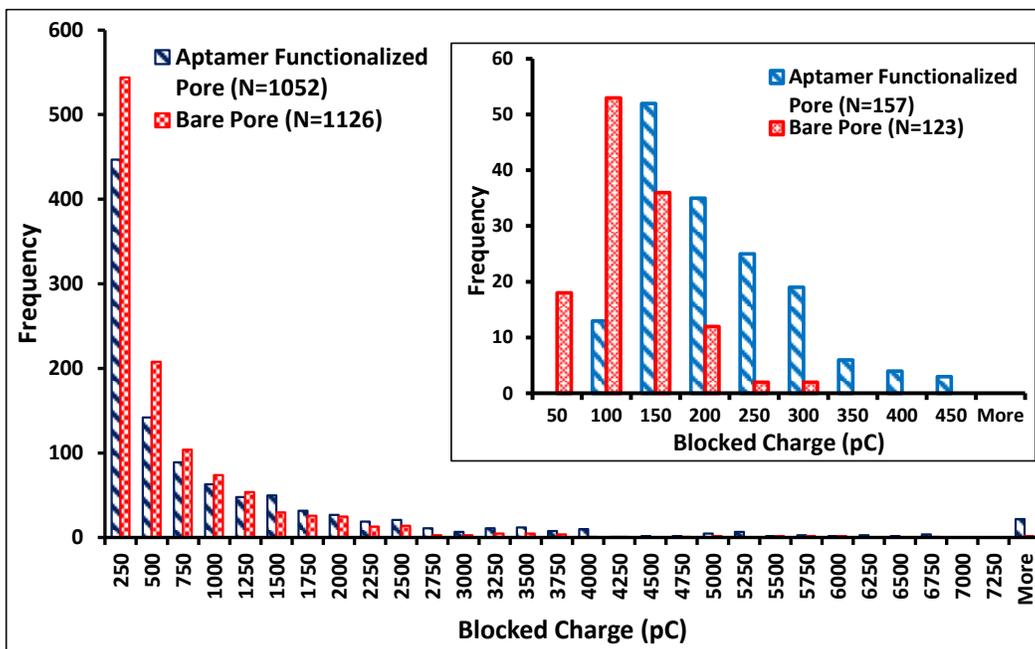


Figure 2-13 Histogram showing distribution of the amount of charge blocked by normal rat blood cells while passing through bare and aptamer functionalized micropores. The inset shows histogram of distribution of amount of charge blocked by hGBM cells while passing through bare and aptamer functionalized micropores [32]

When raw rat blood was run through the micropore under the same experimental conditions, the two distributions of charge blockages were calculated for the bare and aptamer functionalized micropores using this model. The two distributions weren't different as shown in Figure 2-13. Both distributions were single-sided Gaussian in nature and showed peaks on the same ranges of blocked charges. The experimental result

shown in Figure 2-13 can be explained in terms of the interactions between anti-EGFR aptamer and the EGFR present on the surface of hGBM and normal blood cells. As normal blood cells don't have a significant amount of EGFR compared to tumor cells on their surfaces [44], the interactions between the anti-EGFR aptamer and the EGFR on normal rat blood cells was not sufficient to create significant differences between the distributions obtained from aptamer modified micropore and bare micropore. As a result, distributions for blocked charges of normal blood cells obtained from bare and aptamer functionalized micropores were quite similar. Charge blocked by every hGBM cell was also calculated for bare and functionalized micropores. Histograms of blocked charge by hGBM cells are shown in inset of Figure 2-13. The charge blocked while the hGBM cells passed through the pores showed a skewed Gaussian distributions. The maximum occurrence of events for bare micropore was within 50 to 100 pC. On the other hand, the maximum occurrence of events for a micropore which is functionalized with aptamer was within 100 to 150 pC. The comparison between distributions of the charge blocked showed a clear shift of maxima towards the right for aptamer-modified micropore. It implies that due to the interactions between the EGFR on the cell surface and the anti-EGFR aptamer, the amount of charge blocked by the cells was larger. The overexpression of EGFR on hGBM cells caused significantly higher interactions between the anti-EGFR aptamers on the functionalized pore wall and the tumor cells, resulting in the shift of the histograms. This enhanced aptamer-receptor interaction led to spatially separable distributions for bare and functionalized micropores. Despite the overlap, the distribution peaks are clearly located around different maxima. Given the presence of cancer cells in peripheral blood at the outset of cancer [45], the peaks of the histograms can be used as non-invasive diagnostic metrics.

Translocation Behavior

Figure 2-14 shows scatter plots of the pulse widths and current peaks obtained from hGBM cells passing through bare and aptamer functionalized micropores. Although the data points are distributed over a wide region, there is a clear clustering of bare micropore data points within 50 to 450 μs and 0.6 to 2.5 μA (within the dashed boundary).

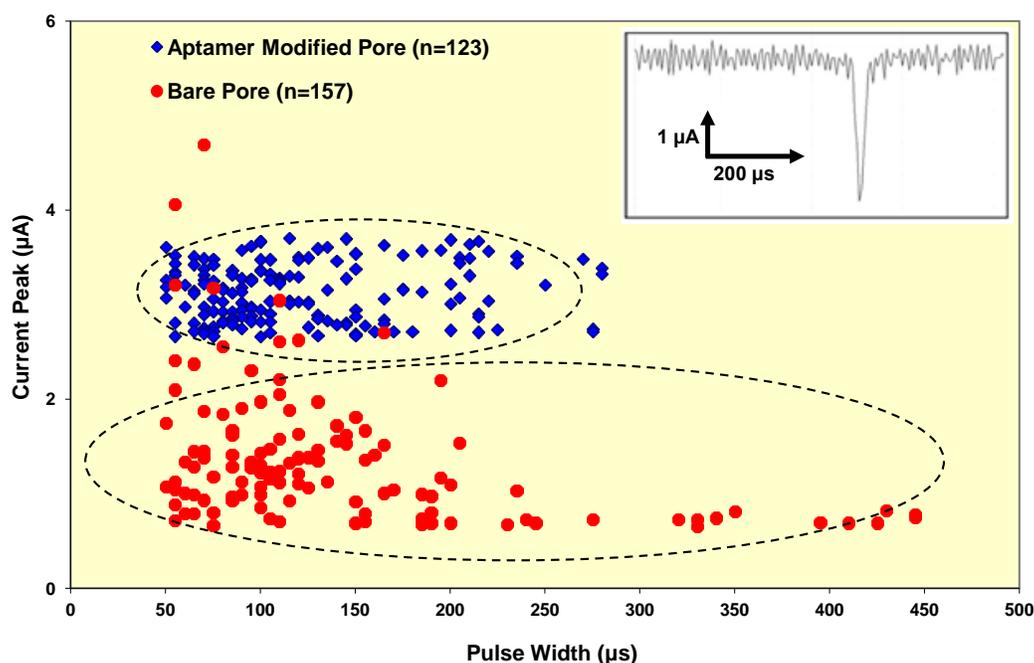


Figure 2-14 Scatter plots of hGBM cells passage through a bare micropore (red circles) and same sized micropore that was functionalized with anti-EGFR aptamer (blue diamonds). The inset shows actual ionic current trace of a typical pulse from an aptamer-functionalized micropore of 20 μm diameter when single hGBM cell passes at a flow rate of 20 $\mu\text{l}/\text{min}$. The electrical data acquisition is done at 200 KHz [32]

The data represented by blue diamonds in Figure 2-14 shows scatter plot of the pulse widths and current peaks for hGBM cells passing through an aptamer grafted micropore. The data points are fairly confined within 50 to 230 μs and 2.5 and 3.8 μA . Even though there is a small overlap but the clustering of data points is spatially separate. The pulses from the aptamer modified micropore clearly have a different center of clustering than that for the bare micropore.

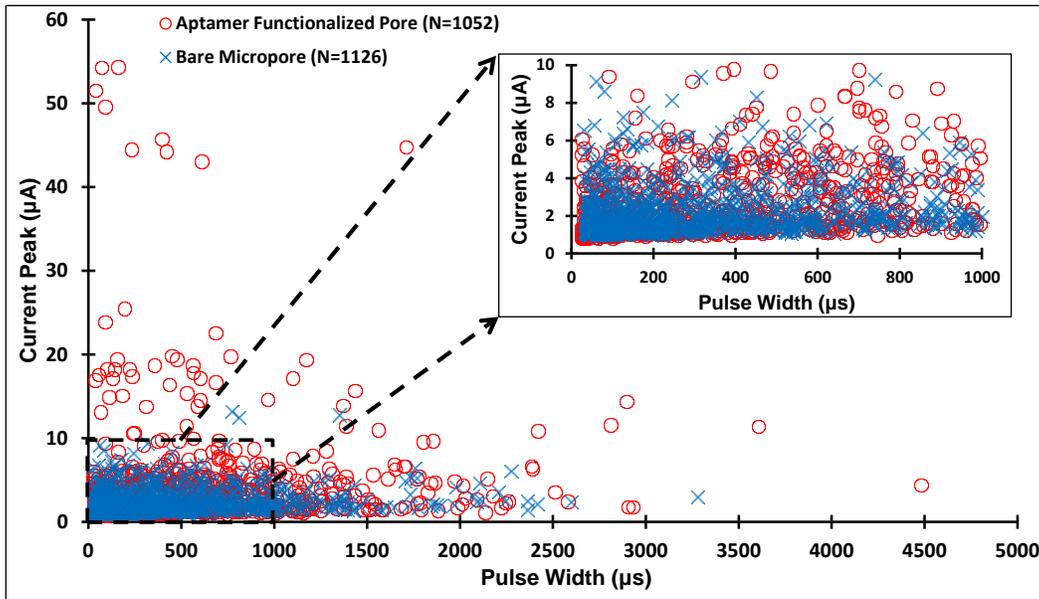


Figure 2-15 Scatter plot of pulse widths versus current peaks of the pulses from rat blood cells passage through a bare (blue cross), and same sized micropore that is functionalized with anti-EGFR aptamer (red circle). Diameter of micropore is 20 μm , flow rate is 20 $\mu\text{l}/\text{min}$. The electrical data acquisition is done at 200 KHz. Inset shows a magnified version of the densely populated region near origin. The functionalized pore data is shown in front of bare micropore data of the main figure. The two data are dense and lie on top of each other [32]

The current pulses were much deeper when the micropore was functionalized with an aptamer. When an hGBM cell passed through a functionalized micropore, the interactions between EGFR on hGBM and an aptamer attached to the micropore's wall reduced the mobility of ions that carried current in the solution. This reduction in mobility led to an effective increase of resistance of the pore. As for the nature of interactions, the most fundamental mechanism here was the binding between the aptamer and EGFR. This binding is still not completely understood as this is neither duplex formation that happens in nucleic acids nor irreversible. The understanding on aptamer-receptor binding hints at formation of secondary structures like hairpins, G-quartets and loops that show selective affinity to certain target molecules. In case of EGFR binding, we have reported before that this interaction is due to the hairpin structure formation in the aptamer molecule, which could be reversed to release EGFR (and the cells) from the surface of the chip [27].

The current blocked by hGBM cells was higher than that of normal cells due to higher resistance, even if their sizes were same (Figure 2-14). As a result, the aptamer functionalized micropore clearly differentiated between cells with more EGFR on their surfaces and cells with less/no EGFR due to its selective interaction with the EGFR on the cell surfaces. On the other hand, Figure 2-15 shows scatter plot of the pulse widths and current peaks obtained from rat blood cells passing through the bare and aptamer functionalized micropores. The data points obtained from both experiments were distributed over a wide region. The pulse width ranges from 50 to 3500 μs and current peak ranges from 0.8 to 15 μA for the bare pore. Similarly, the pulse width varies between 50 to 4500 μs and current peak varies between 0.8 to 55 μA for the aptamer-functionalized micropore. Contrary to the data for tumor cells, the two distributions for blood cells completely overlap. The expression of EGFR on normal cells results in similar

interactions with the bare pore and the aptamer-functionalized micropore. This makes the two distributions spatially inseparable. The experimental set up had the unique ability to differentiate between cells which had a significantly different number of EGFR on their surface. As tumor cells express a sufficiently larger concentration of EGFR compared to their normal counterparts, this unique differentiability can be extended to detect cancer cells from peripheral blood.

Challenges and Solution

One of the major challenges is that some data points associated with cancer cells might be overlapped with the data points of healthy cells due to the wide variation of the levels EGFR expression and sizes. A new methodology was developed to achieve enhanced differentiation. Rather than comparing the data points in a cluster, this new technique determines whether a cell is cancerous or not by comparing the interaction of a single cell with bare micropore and the one modified with aptamer in a single setting.

Single Cell Differentiation

New Experimental Set-up

In contrast to the previous set-up, two micropore chips were sandwiched between three polydimethylsiloxane (PDMS) gaskets. The chips and gaskets were then held together using two Teflon blocks that had holes as shown in Figure 2-16a. Inlets and outlets tubes were connected to the Teflon blocks. Three Ag/AgCl electrodes were required for this set-up. Two electrodes were connected to the inlet and outlet in a similar way as done before and the third electrode was inserted into the PDMS gasket in the middle. The third electrode was pierced through the gasket in a way so that the tip was exposed a little into the hole. The purpose of this little exposure was to make sure when

the ionic solution passed through the hole on the center of PDMS gasket, the electrode would get contact with the solution, thereby completing the circuit. The +5 V and -5 V electrical potential were applied at the electrodes at inlet and outlet respectively. The third electrode was kept at ground potential. The polarities of the applied potentials were chosen carefully to ensure the current flow in the direction of the flow of liquid maintained by the pump. A syringe pump (11 PLUS, Harvard Apparatus) was used to maintain a constant flow rate of cell solution as shown in Figure 2-16a. The cells (5000 cells/ml) were suspended in a 1X PBS and pumped through the system while measuring the current across the two micropores. A typical electrical pulse of an hGBM cell is shown in Figure 2-16b. The process of aptamer functionalization was exactly the same.

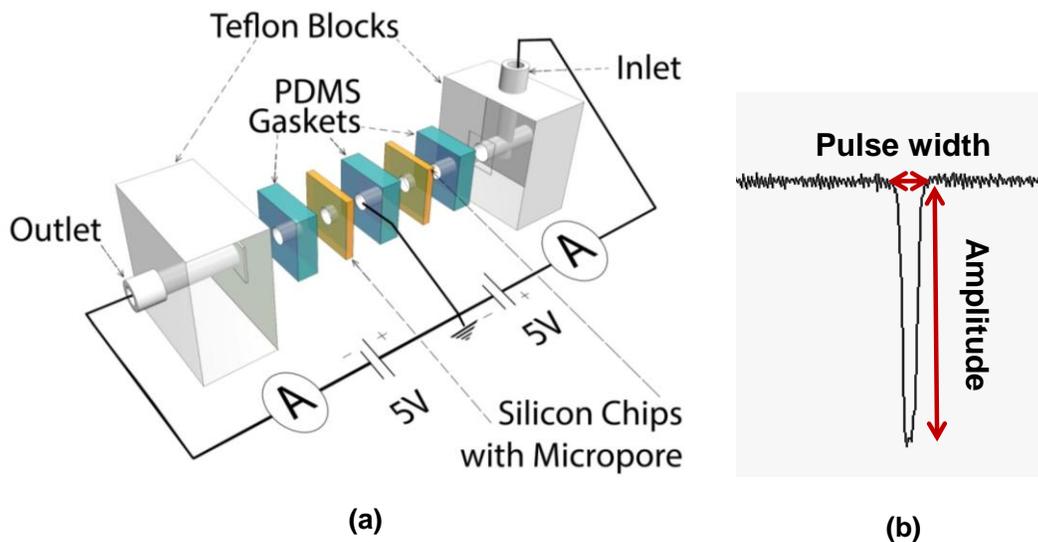


Figure 2-16 (a) Experimental set-up for single cell differentiation using two micropores in series, (b) An exemplary electrical pulse

Results for Single Cell Differentiation

Two identical non-functionalized micropores were used to determine how individual cell behave with each micropore. The electrical pulses obtained from these two micropores were reasonably similar as expected. Figure 2-17 shows the time series of electrical current signal obtained from the first (top) and second (bottom) micropores for hGBM cells. The pulse train from the second micropore was time lagged by approximately 19 ms from the pulse train from the first micropore. This time was taken by each cell travel from the first to second micropore on average.

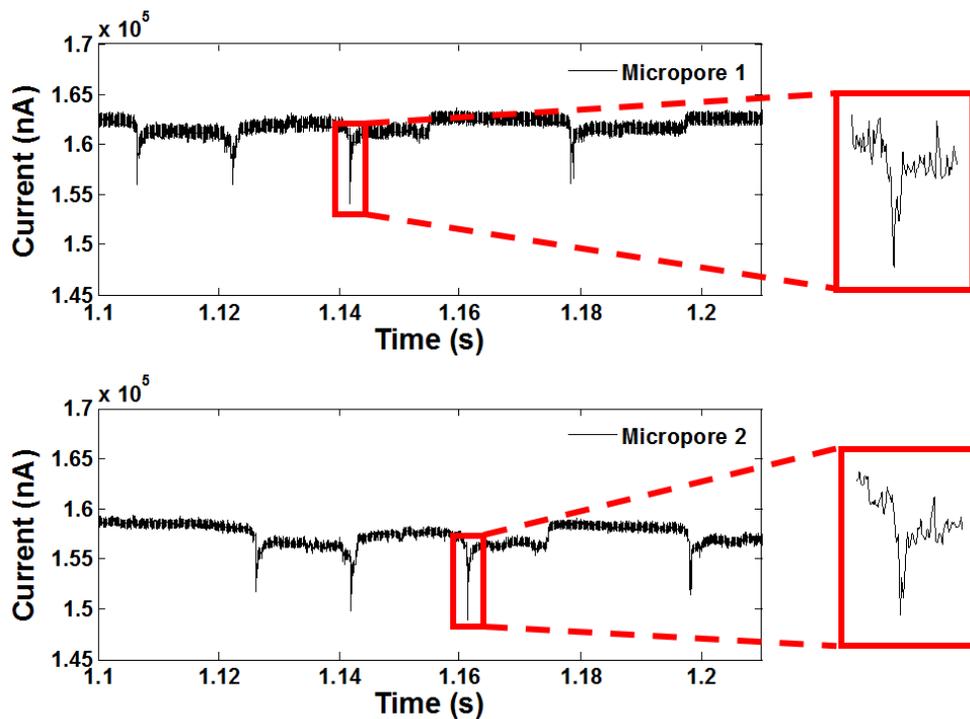


Figure 2-17 hGBM cells detection using two micropores in series

Similar result was obtained for rat blood cells as shown in Figure 2-18. Even though the data was noisier than the hGBM data but still there was one-to-one association between pulses obtained from two micropores. Sometimes the current baseline was wildly shifted due to the presence of various particulates in blood.

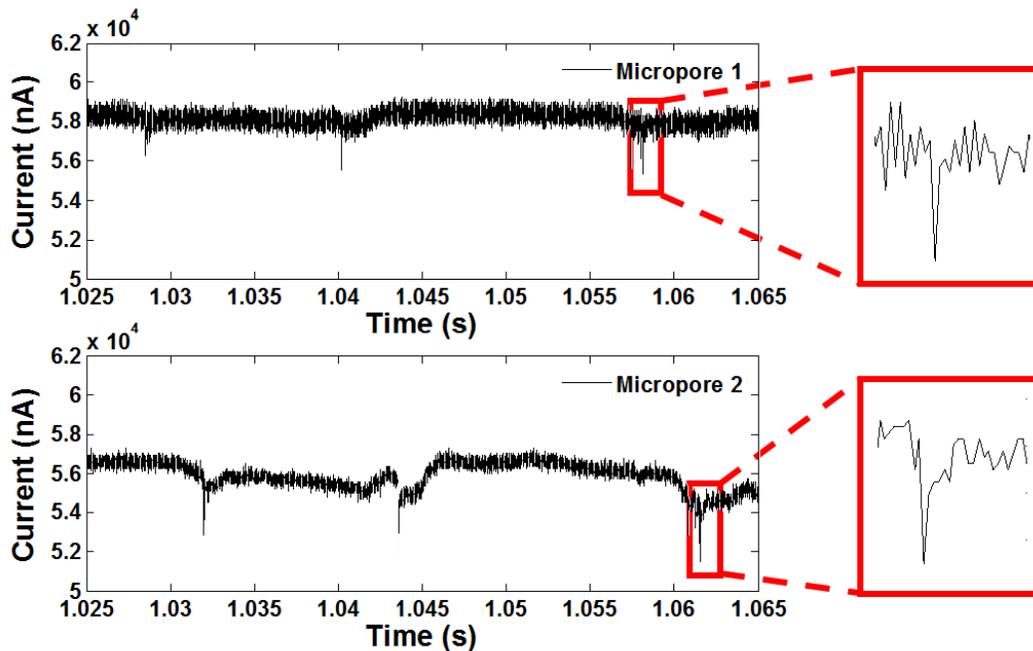
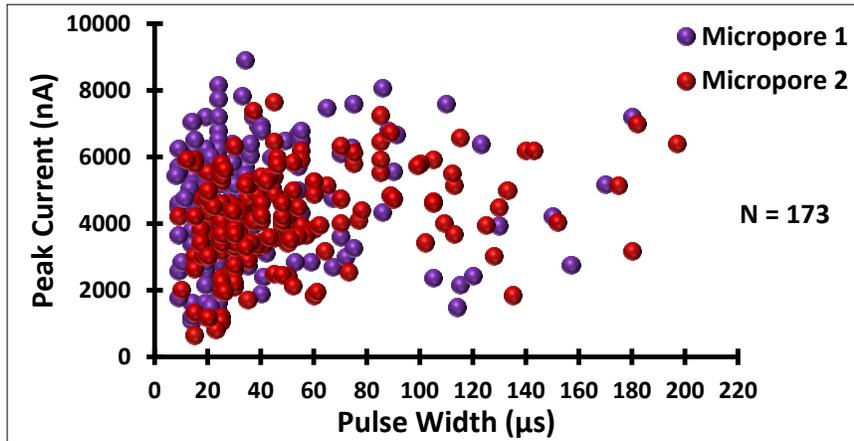


Figure 2-18 Rat blood cells detection using two micropore in series

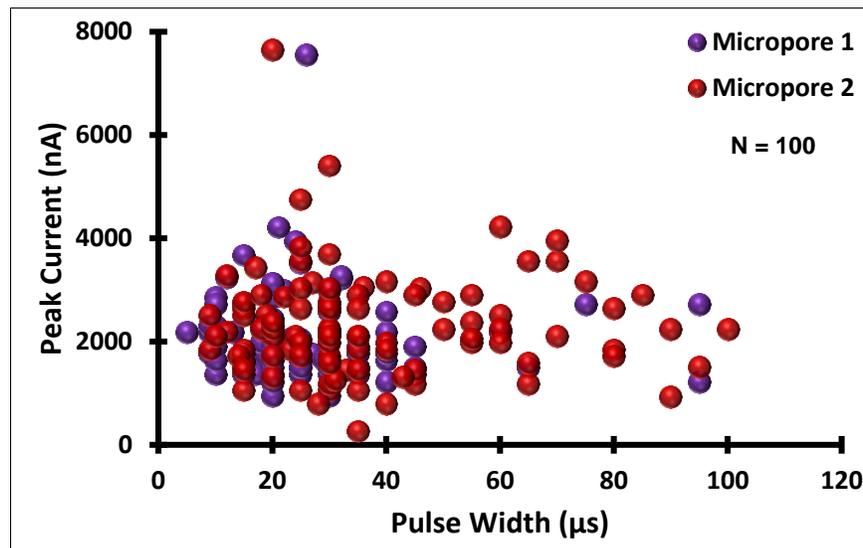
Data points collected from a 10 s window were analyzed and plotted in Figure 2-17. Each data point is represented by pulse width and peak current which were explained in the first section. The distributions of data points obtained from two micropores were fairly similar. The pulse width for micropore 1 and 2 are $38.08 \pm 30.06 \mu\text{s}$ and $50.65 \pm 37.5 \mu\text{s}$ respectively. The peak current obtained by micropore 1 and 2 were $4667.54 \pm 1691.2 \text{ nA}$ and $4210.73 \pm 1399.68 \text{ nA}$. Even though the mean pulse width and peak current were not that close but standard deviation was fairly large i.e. still they could arise from a similar cell when passed through micropore 1 and 2. Most of the time cell maintained a fairly constant travel time from micropore 1 to 2 which was approximately 19 ms. Ideally each pulse should be inverted Gaussian shape, but it took a wide variety of shapes due to the non-uniformity of cells morphology and presence of the noise in the electrical signal. As a result, the peaks didn't appear at the middle of the

signal. It sometimes came before the mid-point of the pulse and sometimes later. To determine the peak position in a pulse, a quantity was computed. If the entire pulse width is considered 100%, at what percent of time the peak appears is called the percent time of peak appearance (PTPA). The PTPA of hGBM cells are 54.03 ± 30.34 and 48.13 ± 19.73 for micropore 1 and 2 respectively. On average, the peaks appeared after 50% of the time of pulse width. Standard deviation was fairly large for this quantity. That means there were wide variety of pulses present in the data pool and some of them were left skewed and some were right skewed pseudo Gaussian in shapes.

For rat blood cells, the pulse width for micropore 1 and 2 were $18.45 \pm 6.64 \mu\text{s}$ and $22.68 \pm 9.35 \mu\text{s}$ respectively. The peak current obtained by micropore 1 and 2 were $2754.29 \pm 1856.25 \text{ nA}$ and $3445.90 \pm 3343.56 \text{ nA}$. The PTPA of rat blood cells were 58.85 ± 17.92 and 51.05 ± 18.43 for micropore 1 and 2 respectively. On average, the peaks of rat blood cells appeared closer to the midpoints of pulses compared to the pulses of hGBM cells.



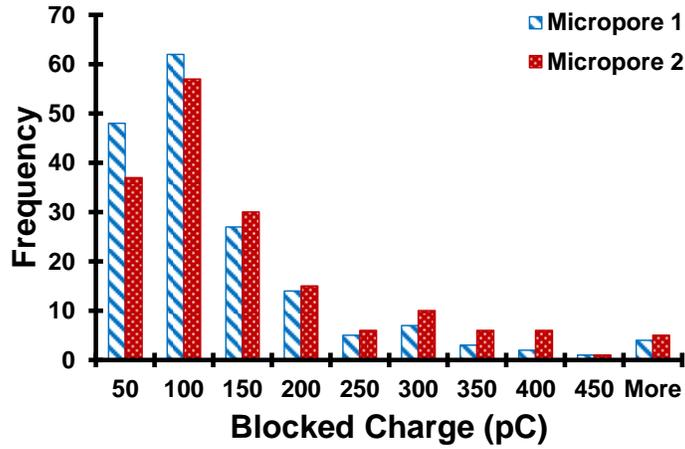
(a)



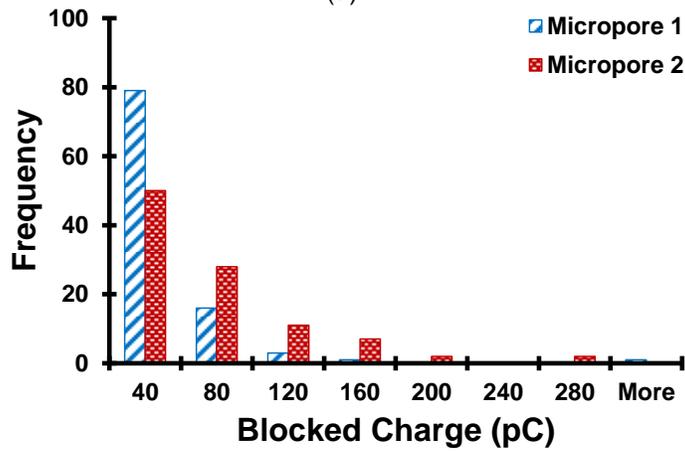
(b)

Figure 2-19 Scatter plot of peak current vs pulse width of (a) hGBM cells, (b) Rat blood cells passed through two micropores in series

The amount of charge blocked by each cell during the passage of micropore is called blocked charge. This is a good parameter to find if there are different kinds of populations in data points. Figure 2-20 shows the frequency distribution of charge blocked by hGBM and rat blood cells while passing through micropore 1 and 2.



(a)



(b)

Figure 2-20 Histogram of blocked charges by (a) hGBM cells (N =173), (b) Rat blood cells (N = 100)

The distribution peaks and shapes obtained from hGBM cells are fairly identical. The distribution of charge blocked by hGBM cell is Gaussian whereas the distribution from rat blood cells are skewed Gaussian in nature. Statistically each cell is interacting with each micropore in a similar way since both of the micropores are in identical conditions. There were four parameters computed to characterize electrical pulses. They were pulse width, peak current, charge blocked by each cell, and PTPA. All of them are

listed in Table 2-2 to compare between the behaviors of cells while passing through two identical micropores.

Table 2-2 Statistics of electrical pulses for hGBM (blue) and rat blood (red) cells (Data collection time 10 sec, N=173, travel time 18.59 ± 2.04 ms)

	Micropore 1	Micropore 2
Pulse Width (μ s)	38.08 ± 30.06	50.65 ± 37.5
Current Peak (nA)	4667.54 ± 1691.2	4210.73 ± 1399.68
Charge Blocked (pC)	107.2 ± 106.81	134.01 ± 126.18
Relative Peak Position (% of pulse width)	54.03 ± 30.34	48.13 ± 19.73
Pulse Width (μ s)	18.45 ± 6.64	22.68 ± 9.35
Current Peak (nA)	2754.29 ± 1856.25	3445.90 ± 3343.56
Charge Blocked (pC)	30.80 ± 25.55	45.95 ± 48.30
Relative Peak Position (% of pulse width)	58.85 ± 17.92	51.05 ± 18.43

Discussion

For the first experimental set-up, data points obtained from the passage of hGBM cells through bare and aptamer modified micropores had different centers of clustering and these were spatially separable. Also, the distributions of blocked charges showed a clear right shift of peaks for pores functionalized with the aptamer. In contrast, normal blood cells didn't show any significant difference in the distributions when passed through bare and aptamer modified micropores. For normal blood cells, both distributions were one-sided Gaussian in nature and there was no relative peak shift observed. These findings were attributed to the interactions between the anti-EGFR aptamer and the EGFR expression levels on the surface of hGBM and normal blood cells. Unlike hGBM cells, normal blood cells don't have overexpressed EGFR on their surface. So their

distributions were not spatially separable as could be done for hGBM cells. This unique capability for analysis can alleviate the challenge of cancer cell data enveloped within normal cell data in a scatter plot as shown in Figure 2-14.

Variation in sizes, amount of charges, receptor-ligand interactions [27], and differences between stiffness of different cell types may induce differences in the translocation behavior [47]. White blood cells (WBCs) have been reported to be in the range of 11-22.5 μm while hGBM cells have been known to be around 20 μm [18]. We used 20 μm diameter micropores. Since the sizes, heterogeneity, and stiffness of cells are similar in normal rat blood, their behavior while passing through the two micropores (bare and functional) was not much different (and hence similar histograms). The inset of Figure 2-13 (hGBM cells) shows the distribution peak is clearly shifted to the right (toward higher blocked charges) for the aptamer functionalized micropore. It implies that with the hGBM cells, with their own size variations, heterogeneity, and stiffness range, the shift in the distribution peak for functionalized micropore came only from the aptamer binding effect. The variations in sizes and other factors can thus be discounted as major contributors in the discrimination effect between normal and cancer cells. The micropore approach has the potential to detect and quantify tumor cells from a blood sample. The same operating principle might be applicable to detect other biological entities such as genes, pathogens, proteins, etc. if their corresponding aptamers are employed. The only challenges will then be to fabricate micropores/nanopores comparable to the sizes of the target entities and functionalize these with the specific ligands.

For the new set-up, it's now possible to compare the interaction of a single cell with a bare and aptamer-functionalized micropore in a single setting. If the one of the micropore is functionalized, then cells with higher EGFR concentration would show different level of interaction with the functionalized micropore from the non-functionalized

one. The interaction can be quantified for an individual cell rather than for a cluster of cells. It can be a new modality to interrogate single cell with two different micropores. The method has the potential to be applied to diagnose other type of cancer or even other disease. The only requirement is to find a target molecule on the cell surface and synthesis of an aptamer for that particular marker.

Chapter 3

Human Voice, a Biometric Signature and a Biospecimen

Introduction

Speech production is a complex process involving coordinated muscular activity. It contains message which is the intended information, and information regarding the speaker conditions of the sound producing apparatuses such as vocal chords, tongue, laryngeal tube, etc. Due to the uniqueness of human voice, it can be used as a biometric signature to identify speaker. It is the process of discriminating one person from others using voice. It has also potential to be a biospecimen to diagnose many diseases of sound producing apparatus. Voice based speaker identification method can be very useful in many areas including security, home or car automation, forensic analysis, tracking, etc. [48]. Many researchers have explored this area with variety of approaches. Elman et.al. studied the hidden features of sound in 1988 [49]. Neural networks have been employed to identify speakers for many years [50, 51]. Reynolds et.al. demonstrated a speaker identification and verification method based on Gaussian mixture speaker model [52, 53]. The process of identifying speakers can be divided into four main steps: voice recording, pre-processing, feature extraction, and decision making. The outline of the method is described in Figure 3-1.

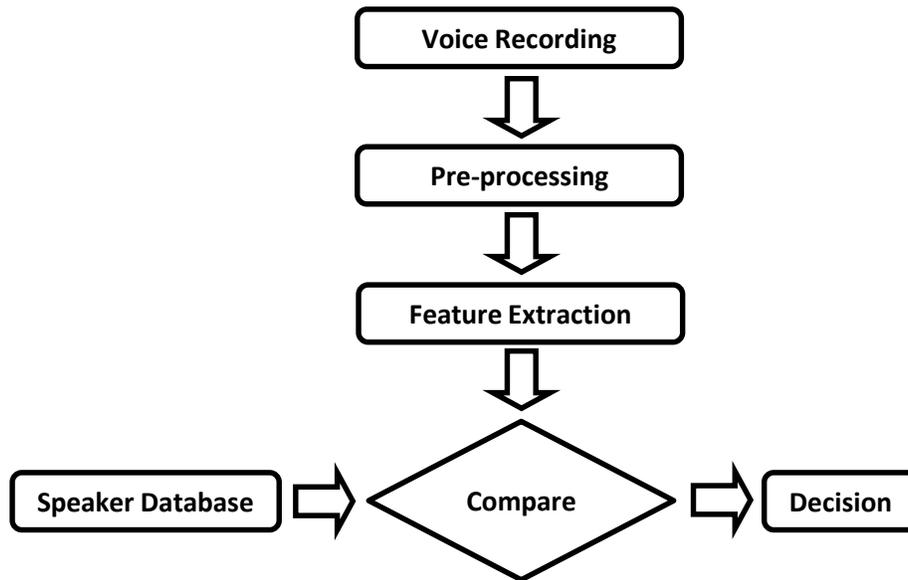


Figure 3-1 Flow chart of a speaker identification algorithm

Voice as a Viable Biometric

A biometric is a mathematical representation of any measurable characteristic of an individual. Due to sheer number of possible variations in a speaker's voice, it is extremely difficult to mimic someone's voice for biometric authentication. Our ears may not perceive the subtle differences in the voice of an imposter, but an authentication system can computationally analyze multiple features ranging from spectral magnitudes, formant frequencies, time domain features, pitch, and probably phase spectrum as well. As a result, differences that are not discernible to our ears can be caught in a voice based authentication system. Similarly, recording of a voice may sound like a perfect match to ear, but detailed mathematical analysis reveals many underlying features (distortion, traits of the recording system, etc.) that set apart the actual speaker from the imposter. An even more robust method would be to employ a text independent voice authentication that uses a randomly generated text to prevent the forger from knowing it

ahead of time. The probability of a fraud to successfully pass through such an authentication system is exceedingly small.

Reliability of Voice Based Authentication System

The primary task of any biometric system is to allow a user to access a system with less number of trials and at the same time prevent any fraudulent access to someone else's account. The latter is actually a more stringent feature of a reliable authentication system. No authentication system based on biometrics is perfect. There is always a tradeoff between the limit on number of false negatives and false positives.

A common feature of all biometric based authentication systems is that all of them are probabilistic in nature to some extent. As a consequence, there is always a little chance of producing wrong result sometimes. When a biometric is recorded, there are always some artifacts in the recording, digitization, and encoding processes. Robustness of a system is defined by how small these errors are. For any biometric authentication system, a threshold probability level is set above (or below) which a new sample is accepted (or rejected) when compared to the stored metric. It is possible to control the number of false positives and false negatives by changing the threshold. The system can tolerate more false positives or negatives depending on particular application. As a result of multi-feature authentication, voice biometric usually offers better authentication than other biometrics.

Over the last few years, human voice has become a viable biometric like a fingerprint or iris scan. Voice offers a distinct advantage over other biometrics. Contrary to other biometric authentications, voice based authentication doesn't require the subject to be present on the spot where the authentication is done. Additionally, voice based authentication system has the ability to accept secondary information from the subject in

parallel with the authentication process. For example, a new pass code can be input while the voice biometric gets processed. This capability of taking additional information in the authentication process offers a unique advantage over other systems in real world applications.

A simple text-dependent speaker identification and verification system has been developed that identifies (verifies) a speaker with a mere utterance of vowels. The system records vowels from a subject, processes it, and extracts few features to be used later to identify subjects. A database was formed with the extracted features first. Magnitude spectrum of fast Fourier transform (FFT) was used as the feature in the analysis. The authentication was done by finding the percentage match of the sample's feature with that stored in the database.

Background

Sound is a mechanical wave that is produced by any vibrating object. The vibration then propagates through a medium (air, water, metal, etc.). The source of disturbance could be the vocal cords of a person, a tuning fork, a vibrating string, sound board of a guitar or violin, diaphragm of a speaker, etc. All the vibration sources create a back and forth motion in the propagation medium. The number of complete cycle of back and forth per unit time is called the frequency of the sound. A commonly used unit of frequency is Hertz (Hz). If an object makes 100 complete cycle of vibration in 2 seconds, its frequency is 50 Hz.

Our daily communication is almost entirely dominated by speech. It is the primary source of raw information used by a listener to recover the message. A sound wave is the end product of a complex speech production mechanism. It is a disturbance that results from vibration and propagates through any elastic medium. It can be produced from

anything that vibrates in audible range (roughly 20 Hz to 20 KHz for humans). A series of coordinated muscle actions is required to produce a sound. The whole process of speech production is controlled by our brain. A cross-section of larynx is shown in Figure 3-2. In the beginning of sound production, the vocal cords are closed. The air pressure from the lung keeps building due to this closure. When the muscle of vocal cords can't hold the pressure any more, it suddenly lets go the air in the form of little pops, buzzes, and hisses. Then, these sounds are filtered in the laryngeal tube and modified with the help of the lips, tongue, soft palate, jaw and other articulators to produce meaningful sounds [54].

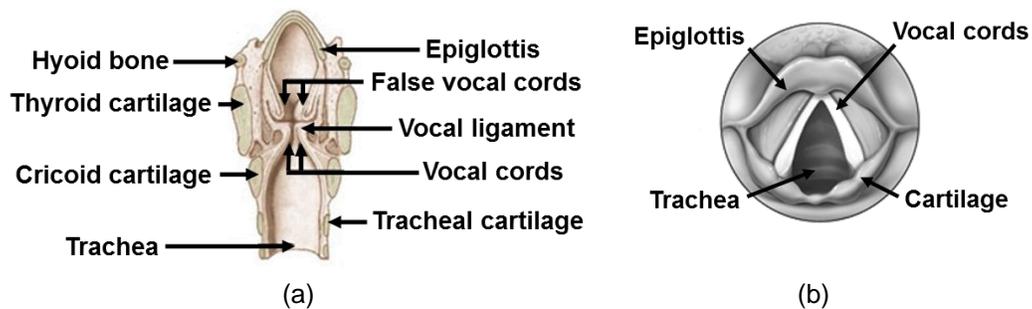


Figure 3-2 Human larynx (a) Vertical cross-section, (b) Horizontal cross-section (adapted from [55, 56])

The rate at which the vocal cords chop the air flow is called the fundamental frequency. When the sound passes through the laryngeal tube, it resonates depending on the shape of the tube. Each resonance will produce a spectral peak in the sound spectrum. These spectral peaks are known as formant frequencies. There are few fundamental and formant frequencies associated with each vowel. For any vowel, first three formant frequencies are most important characteristics of any vowel. The range of frequencies used to pronounce a vowel may vary depending on the speaker. As a result, distributions of frequencies in the sound spectrum can be used as a metric to identify speaker. It may also be possible to find a trend in somebody's voice by monitoring the

spectrum over a period of time. It might tell if there is a consistent change in the vocal chords. In case of paralysis of vocal cords, or formation of lump, polyps, or soars on the vocal cords, the frequencies may change. If a database of magnitude spectrum is formed under healthy conditions, it can be used as a reference to detect anomaly in the voice of other subjects noninvasively. Since the larynx acts as one end closed tube, it produces only the odd harmonics i.e. the formant frequencies are odd multiple (1, 3, 5, etc.) of the fundamental frequencies. If there is a little lump formed on the vocal cord, it will change the fundamental frequency. A little change in fundamental frequency will be multiplied by 3 or 5 for second or third formant frequencies respectively. Due to this inherent frequency multiplication, it might be possible to detect anomaly on vocal cord at very early stage of a disease by observing any consistent frequency shift in the spectrum. Since vocal cords act like a variable length tuning fork made up with material of variable elasticity, it can produce a wide range of frequencies by changing the muscle tension. If a little mass is added on a leg (wax) of a tuning fork, its fundamental frequency changes. Similarly, any nodule, polyp, bulge on the vocal chords would change the fundamental frequencies as well as the formants of a vowel.

Our ears receive this complex sound and they break into its frequency components in a similar fashion as the prism splits white light into color components [57-59]. If the ear couldn't do that, it wouldn't be possible to recognize people just by hearing their voices. In principle, an individual speaker can make a finite set of sounds in almost an infinite number of ways in time domain. It is impossible for human brain to store all these time domain permutations of sounds.

In order to recognize a person, the brain must find a match between what it hears and what is stored in it. Since the time domain information is too much to remember, the brain must store the frequency domain information to identify speakers [60]. Since our

brains can identify subjects using the frequency domain information, similar method was adopted for this study to emulate the identification.

Voice disorder is a relatively common pathological condition affecting the sound producing apparatus in human. Based on the sources, voice disorder can be classified as physical, neuromuscular, traumatic, and psychogenic. All of these distort the quality of the voice. Like any pathological condition, voice disorders, also need to be diagnosed first before treatment. The most common tools for voice pathology are videostroboscopy [61] and videokymography [62]. However, these current methods are invasive in nature, expensive, and time consuming. A noninvasive and simple method would be ideal for diagnosing voice disorder.

Methods

Data Collection and Database Formation

Primarily, three subjects voluntarily participated in developing an audio sample database. Five distinct English vowel sounds (A, E, I, O, and U) were chosen for each individual subject to speak. Each audio sample was recorded using a standard computer for the duration of 2 seconds with 22050 samples per second. MATLAB was used to control the recordings of the sounds. Five samples of each sound from each subject were recorded. Fast Fourier transform (FFT) was performed to calculate the magnitude spectra of the samples. The average of these magnitude spectra was considered as the signature of a particular individual for that sound. It was stored to form the database. Five samples were recorded to retain the consistency and reliability of the collected samples.

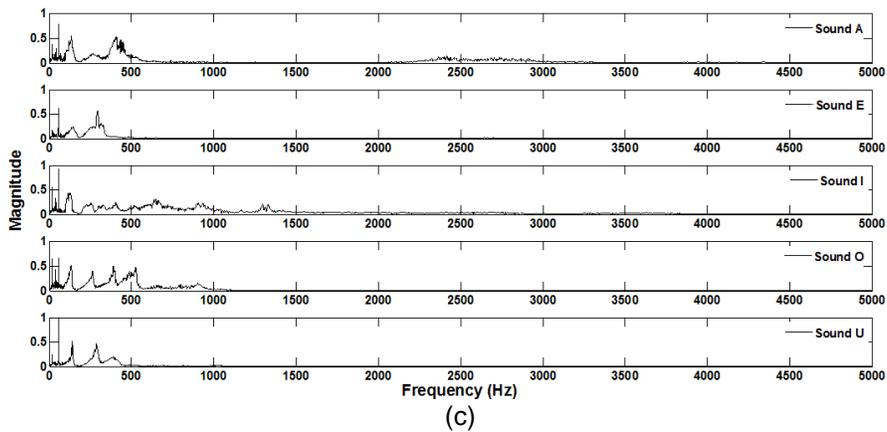
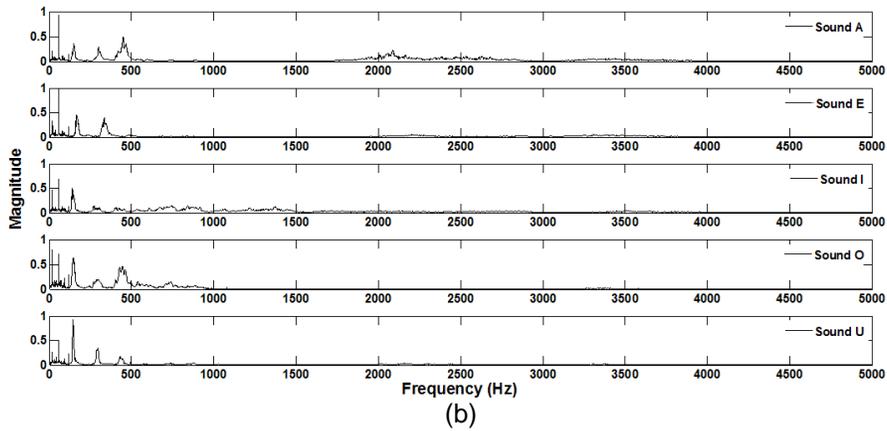
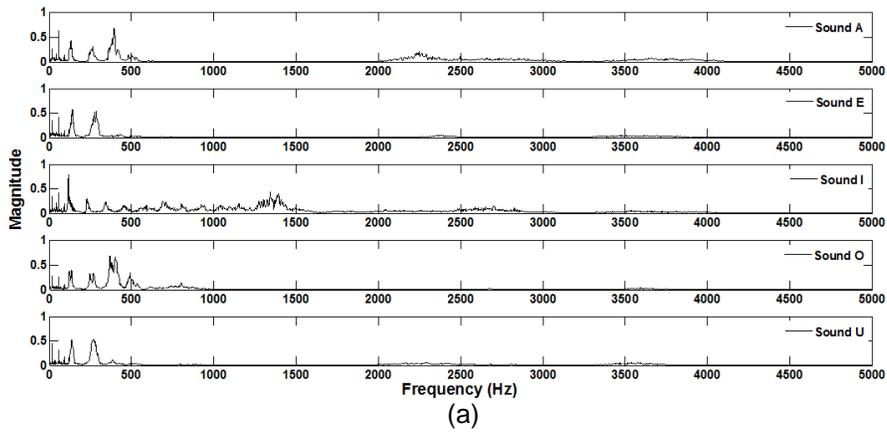


Figure 3-3 Average of normalized magnitude spectra of vowel sounds for (a) Subject 1, (b) Subject 2, and c) Subject 3 ($n=5$ for each vowel). Partial spectrum (0 – 5 KHz) is shown since most of the features lie within this region

Feature Extraction

Once recorded, FFT was performed on each audio sample to analyze the frequency domain information. The average of magnitude spectra was calculated from every dataset of five samples and served as signature template for future authentication. It was the feature that was utilized to identify subjects. A well-defined but distinguishable pattern was observed among different signature templates. A time domain and its corresponding frequency domain representation of a sample sound (Sound of "A") are shown in Figure 3-4. The distinctive frequency components are clearly visible from the magnitude spectrum in Figure 3-4b. Most of the information is confined within 0 to 5 KHz.

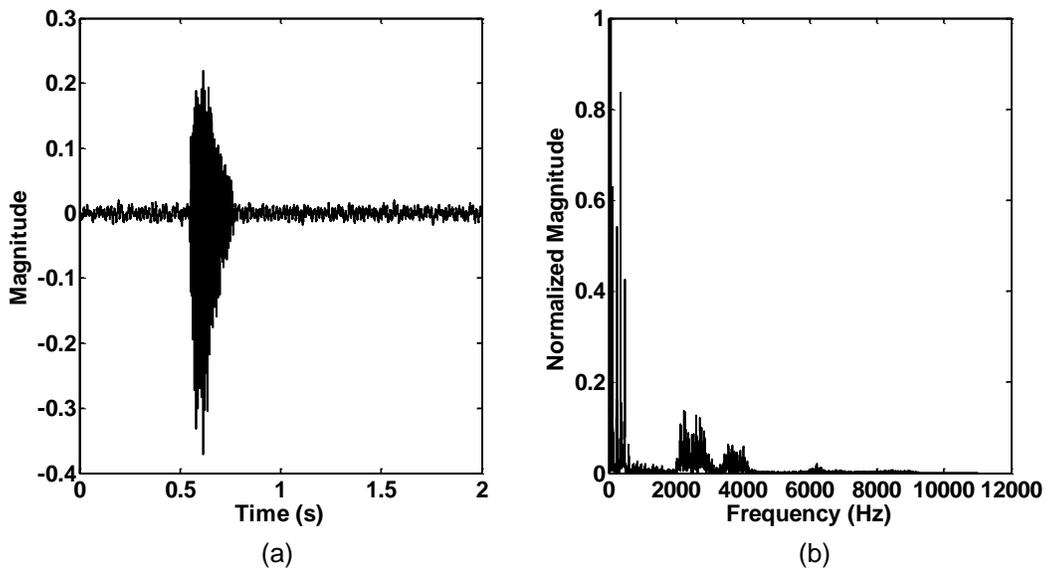


Figure 3-4 (a) Time domain representation of Sound A (recording time = 2 sec, sampling frequency = 22050 Hz), (b) Normalized magnitude spectrum of that signal

System Verification

To verify the system, another set of samples of each kind was collected from all subjects and cross-referenced with the available templates from the database. The

standard correlation coefficient was calculated in order to find matching between sample and stored template. The correlation coefficient represented the percentage match of the samples with templates. The magnitude spectrum of each sound pronounced by different subjects was compared graphically and mathematically. The results provided substantial support to the hypothesis that a subject could be identified and verified using their voice as a biometric. It only requires computing the correlation coefficient between the magnitude spectrum of a sample sound and template stored in the database, then to check if the coefficient passed a threshold or not. Setting up the threshold was dependent on the target application. In our study, 70% matching was used as threshold. Any match above (below) 70% was considered as positive (negative) hit.

Results

Algorithm 1

From the Figure 3-4b, it is evident that the magnitude spectrum of a particular sound was divided into regions that contained most of its features. The magnitude spectrum contained information about the type of sound, speaker, pace, tone, pitch etc. Any sound was the final output function of these variables. Variation in any variable changed the output depending on how strongly the outcome depended on that particular variable. The information about the speaker was one of the many aspects of the magnitude spectrum. The multi-features nature of speech makes the speaker's identity concealed. As a result, it was really difficult to extract the speaker's identity from this blend. Any variation of pace, tone, pitch, mood, etc. would make the situation even worse. For an effective identification, it was very crucial to keep all the unwanted variations at minimum level. To ensure this, five samples of each sound (A, E, I, O, and U) were taken under the same environmental conditions from each subject, with the

same experimental set-up, at the same time of the day. The configuration of the data acquisition system was kept exactly same for all cases. All the subjects were fairly calm and quiet during the sample collection. All samples were collected under the consent of the speakers. Once the sample collection was done, all the samples were converted into frequency domain by computing FFT. The average of FFT magnitude spectra was then calculated. The magnitude spectra were used as the signatures of the subjects. Averages of magnitude spectra of A, E, I, O, and U sounds of Subject 1 are shown in Figure 3-3a. The magnitude spectrum shows distinct characteristics of the sound for each speaker. Each spectrum from a single speaker is significantly distinct from others. The variations are clearly noticeable in the plot. Similarly Figure 3-3b and Figure 3-3c are the averages FFT magnitude plot of Subject 2 and 3, respectively.

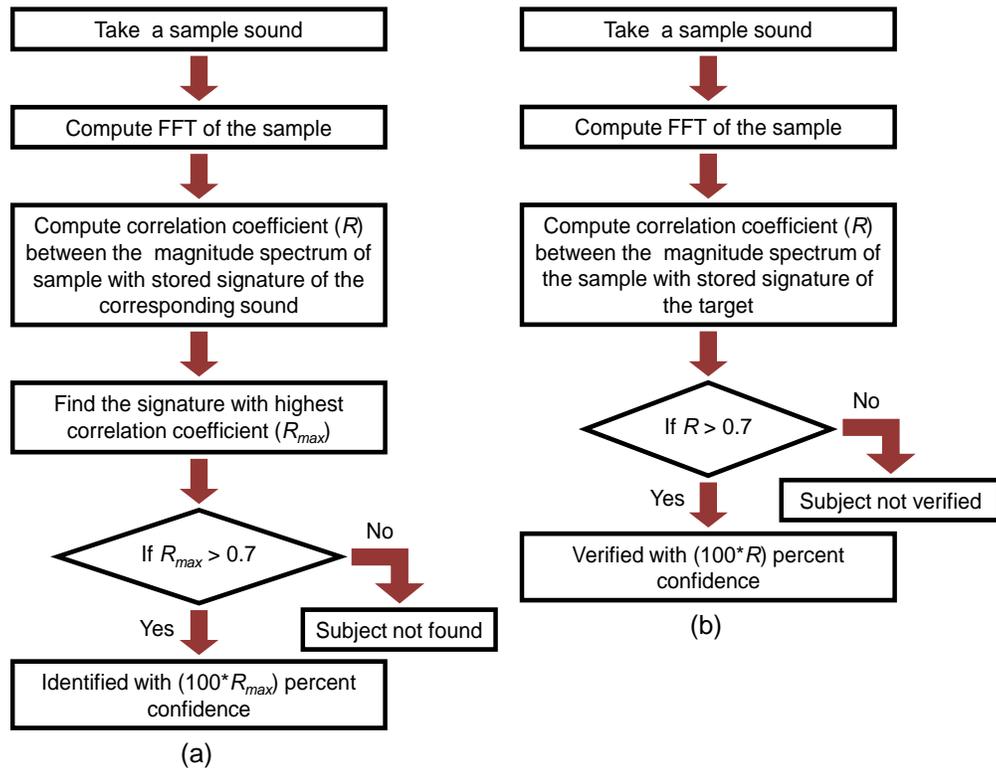


Figure 3-5 Flow chart of Algorithm 1 (a) Identification, (b) Verification

The flow charts of the speaker identification and verification processes of Algorithm 1 are shown in Figure 3-5. Figure 3-6 shows the resemblances in FFTs of particular vowels spoken by three speakers. Different sounds had their own characteristic frequencies. Different individuals produced the same sound with the emphasis on different frequencies. The combination of these two phenomena gave rise to unique signatures of each sound made by an individual. From the Figure 3-6, it seems that the low frequency region (below 250 Hz) is quite similar for all subjects. On the other hand, the frequency region of 250 - 3500 Hz shows significant variations among different speakers.

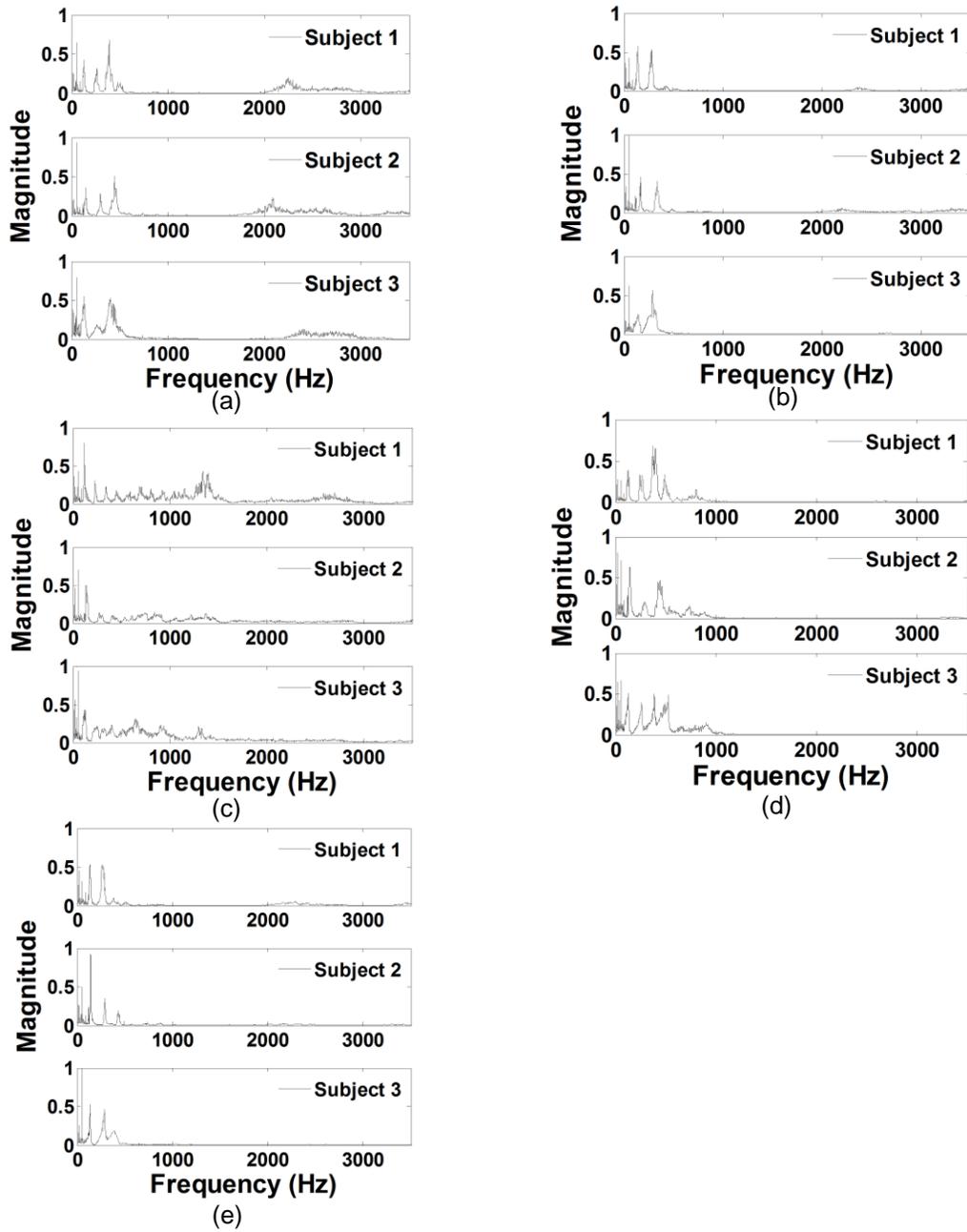


Figure 3-6 Average normalized magnitude spectra (0 – 3.5 KHz) of (a) Sound “A”, (b) Sound “E”, (c) Sound “I”, (d) Sound “O”, and (e) Sound “U” of three subjects (n=5 for each average)

Table 3-1 Correlation coefficient between the magnitude spectra of the test samples and signatures

		Signature 1					Signature 2					Signature 3				
		A	E	I	O	U	A	E	I	O	U	A	E	I	O	U
Test Sample of Subject 1	A	0.88					0.36					0.73				
	E		0.78					0.14					0.53			
	I			0.75					0.57					0.64		
	O				0.77					0.38					0.60	
	U					0.68					0.15					0.44
Test Sample of Subject 2	A	0.23					0.70					0.35				
	E		0.22					0.70					0.51			
	I			0.40					0.79					0.48		
	O				0.31					0.86					0.38	
	U					0.27					0.57					0.31
Test Sample of Subject 3	A	0.70					0.52					0.90				
	E		0.85					0.23					0.76			
	I			0.65					0.60					0.86		
	O				0.74					0.53					0.91	
	U					0.83					0.40					0.91

To find the percentage matching of test sample and the stored signatures, we computed standard correlation coefficient. The correlation coefficients are charted in Table 3-1. Each row represents a test sample from a subject and each column corresponds to a stored signature. Table 3-1 is visually presented in Figure 3-7. Each sound from each speaker showed the highest match with its corresponding signature in all, but one. The technique was tested for three subjects using five vowel sounds (total 15 cases). Only one out of 15 cases (circled) failed to produce highest correlation with its own signature. Sound "E" from Subject 3 produced 85% match with the signature of sound "E" of Subject 1 and 76% match with the signature of its own. It should have had maximum matching with its own signature, instead of others. As a result, the accuracy of the technique was 93.33%. There are multiple reasons for failing to produce the highest correlation with its own stored signature. Differences in the feature extracted from sounds recorded at different times of the day (morning, noon, night) may be responsible for this. The feature can also change depending on the mood of the person. Aging and sickness might also cause some variations. Moreover, variability in pitch, pace, and tone introduce enough variations to produce false results. Environmental noise can play a crucial role in the identification process as well.

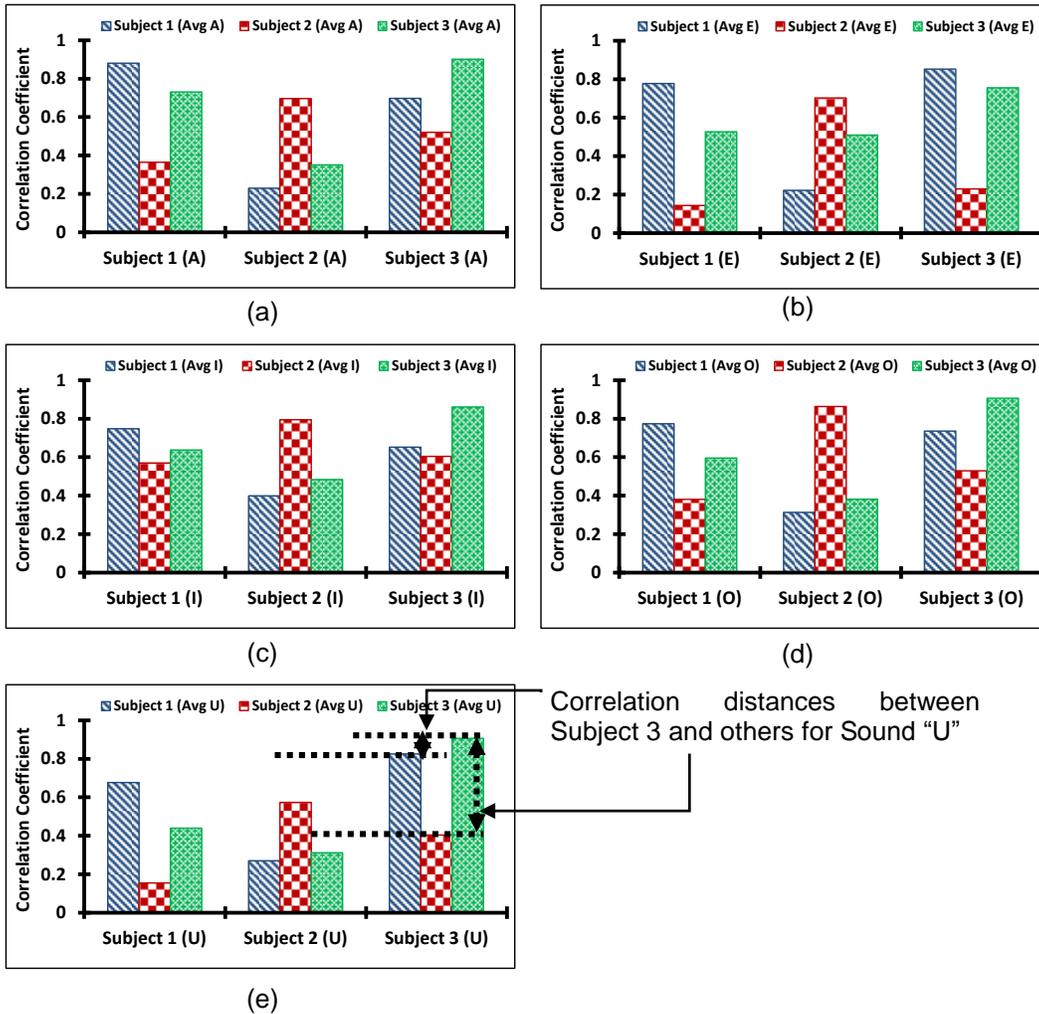


Figure 3-7 Correlation coefficient between the magnitude spectrum of test samples and their corresponding signatures for three subjects. X-axis represents test samples of a vowel pronounced by three subjects; Y-axis represents the correlation coefficient, (a) Sound "A", (b) Sound "E", (c) Sound "I", (d) Sound "O", and (e) Sound "U"

When a test sample was cross-referenced with a number of stored signatures including its own, the matching was the highest most of the time, with its own signature.

The correlation coefficients with other signatures were not negligible though. For some cases, a test sample showed significant amount of matching with others' signatures.

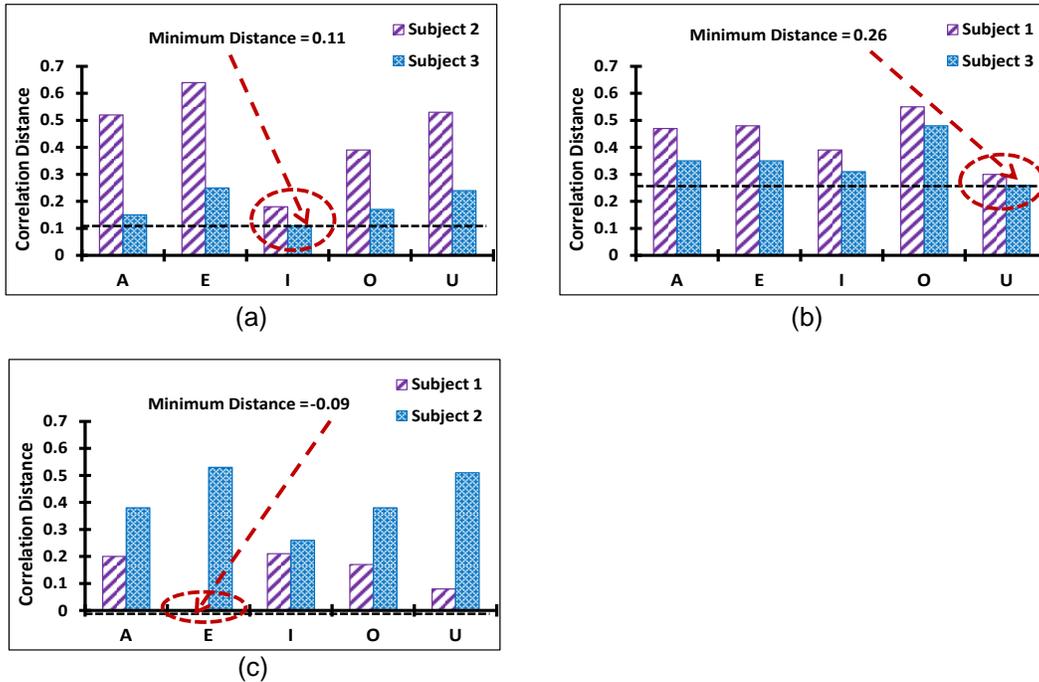


Figure 3-8 Correlation distance between test sample of a subject with other two subjects, (a) Subject 1, (b) Subject 2, and (c) Subject 3

The significant matching with others makes it difficult to decide whether the sampled person is the person he is supposed to be. The number by which the correlation coefficient between sample and its own signature differs from that with others gives a margin on how well a subject can be differentiated from others. To quantify this margin, we defined a metric, called "correlation distance". It is defined as the difference between the matching of a sample with its own signature and that with others'. Figure 3-7Figure 3-7e explains graphically what correlation distance is. For any identification and verification process, the more correlation distance it would produce, the easier it would be for the system to identify target from others. To analyze our technique, correlation

distances for all cases were calculated. The results are shown in Figure 3-8. Correlation distances of Subject 1, 2, and 3 are plotted in Figure 3-8a, b, and c, respectively. The amount of memory storage required to store signatures of 1000 users is 168.23 MB. The identification and verification times using this technique were 557 ± 23 and 1.762 ± 0.151 ms, respectively.

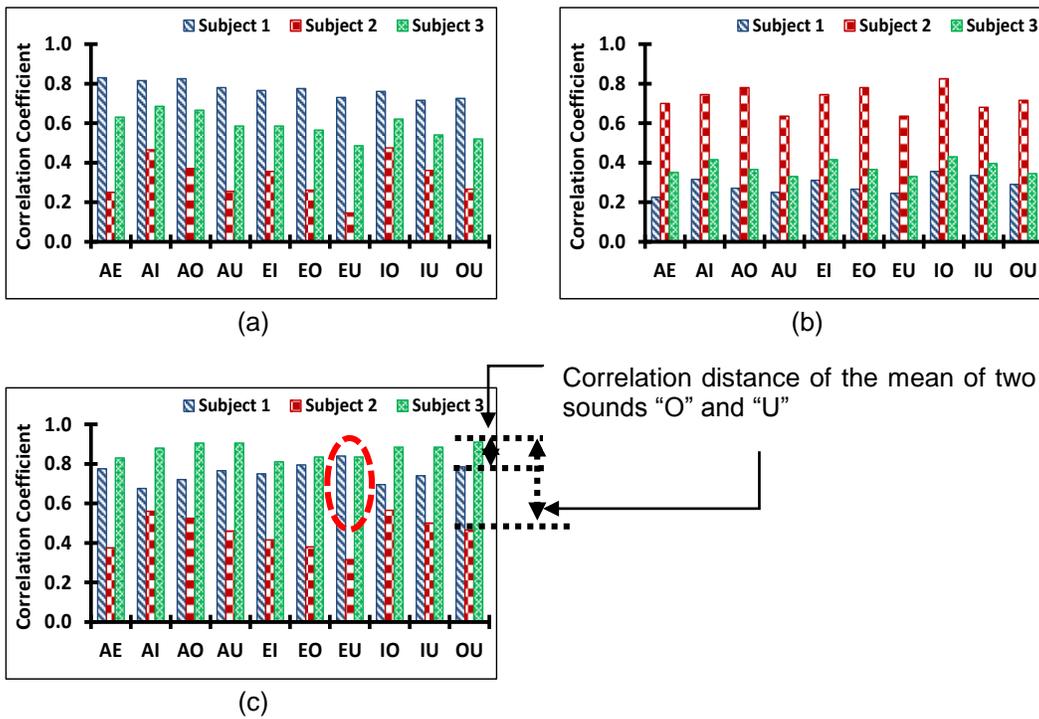


Figure 3-9 Mean correlation coefficient vs pair of sounds. X-axis represents the pair of sample sounds for a particular subject used to find the mean; Y-axis represents the mean of these two correlation coefficients, (a) Subject 1, (b) Subject 2, and (c) Subject 3. Part c also explains correlation distance

Algorithm 2

In order to improve accuracy of authentication of Algorithm 1, a new technique was developed. The flow chart of this new algorithm is shown in Figure 3-11. In this

technique, we computed FFT of two different sounds (for example “A” and “E”) instead of one. First, correlation coefficients were calculated for each sample sound with the stored signatures of the same sounds of all subjects. Then the mean of correlation coefficients of two sounds of each pair from the same subject was calculated. It is possible to form ten unique pairs with five vowels (AE, AI, AO, AU, EI, EO, EU, IO, IU, and OU). The mean of each pair was then used to decide whether the target was found or not. Figure 3-9 shows the average correlation coefficients for each pair of sounds of each subject. Correlation distances as shown in Figure 3-10 were computed to compare the performances.

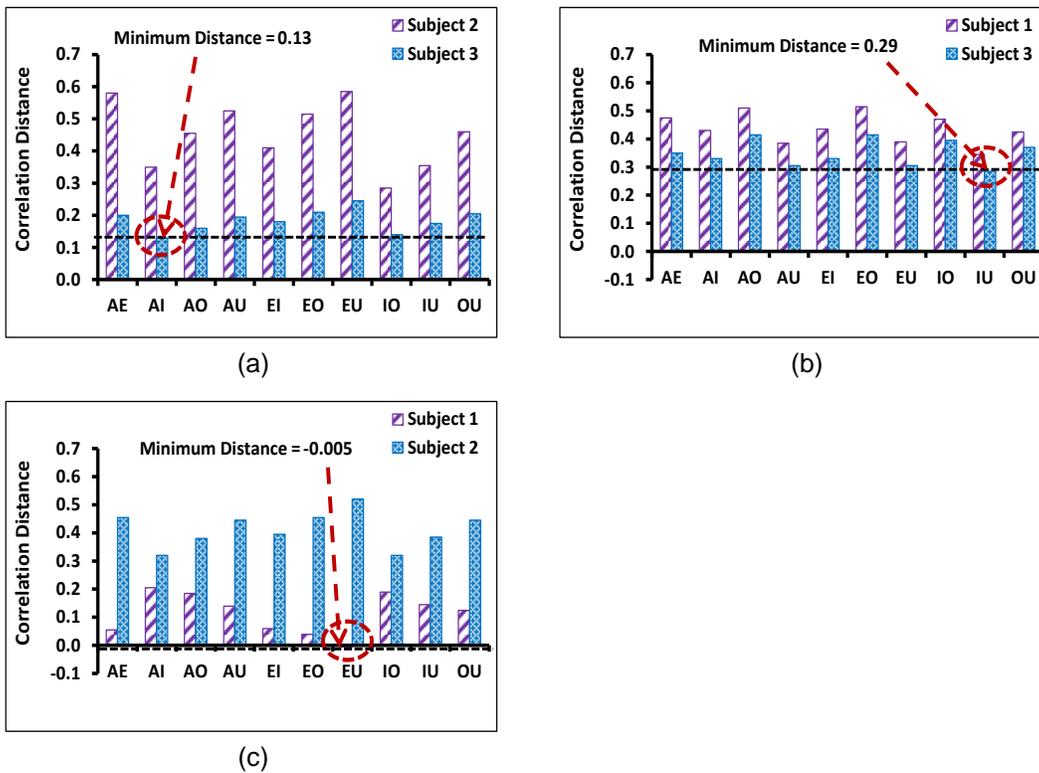


Figure 3-10 Mean correlation distance for pair of sounds of one subject with that of others

(a) Subject 1, (b) Subject 2, and (c) Subject 3

Correlation coefficients and correlation distances were used as two metrics to compare the techniques. It was observed that the minimum correlation distances for all subjects (three) improved without sacrificing much of the other metric, correlation coefficient. The percentage increase of minimum correlation distances for Algorithm 2 with respect to Algorithm 1 was 18.18, 9.62, and 94.44 for Subject 1, 2, and 3, respectively. Five different sounds can form ten distinguishable groups, each with three sounds. The performance metrics for groups of triplet sounds i.e. FFT of three separate sounds (for example A, E, O) were also taken to calculate the average. The performance in terms of correlation distance improved further while the correlation coefficient didn't change much. In the case of group of three sounds, 27.27%, 19.23%, and 155.56% increment of minimum correlation distance for Subject 1, 2, and 3, respectively were observed compared to Algorithm 1. From the above analysis, we have found that the margin (correlation distance) of decision making proportionately improved with the number of sounds taken per group to calculate the mean. Mean of multiple correlation coefficients worked better than the correlation coefficient of a single sound.

But there was a tradeoff for this improved performance. The identification and verification time increased depending on the number of sounds per group used to calculate the mean. Additionally, the storage space was proportional to the number of signatures needed to be stored per person. If two separate sounds were needed for identification, two signatures per person had to be stored. This doubled the storage requirement. For N number of signatures per person, it would be N times of the storage of a single sound. Out of 30 cases (10 pairs of sound from each subject), it failed to produce highest match for one case as shown in Figure 3-9e (red circled). The accuracy of the method was thus 96.67%. Algorithm 2 offered better accuracy than Algorithm 1, but it was slower. The identification and verification times were 1126 ± 41 and 3532 ± 261

ms, respectively. The memory requirement for Algorithm 2 was 336.46 MB per 1000 users.

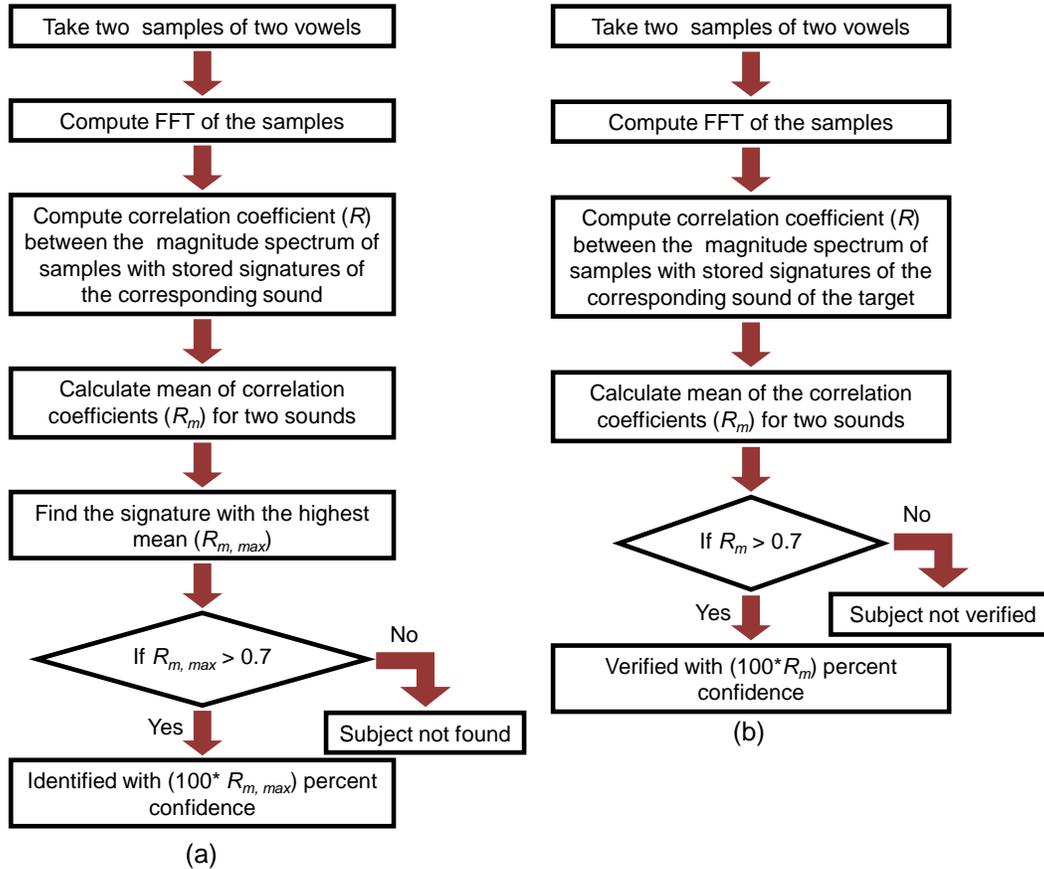


Figure 3-11 Flow chart of Algorithm 2 (a) Identification, (b) Verification

Algorithm 3

In order to reduce the identification time of Algorithm 2, a third Algorithm was developed. The flow chart of the Algorithm 3 is shown in Figure 3-12 and Figure 3-13. It had all the advantages of Algorithm 2, but accelerated the process of decision making. Once sample of a sound (for example, Sound “A”) was taken from a subject, its FFT was computed, and correlation coefficients were calculated against all stored signatures of the sound (Sound “A”). Maximum of all coefficients was obtained. All cases with less than

50% match were filtered out. It eliminated most of the signatures that were not possible matches. Only a fraction of cases remained. These remaining signatures went to the next round of elimination. Due to the reduction in the number of signatures that were potential match, the decision making became faster. In the second phase of elimination, one sample of second sound (for example, Sound "E") was taken. Correlation coefficients were also calculated against all signatures of this second sound (Sound "E") for the rest of the subjects. All signatures that failed to produce more than 50% match with the sample were eliminated. Mean of the two coefficients produced for first and second sound ("A" and "E") for each signature was computed. After the second round of elimination, the signature that produced highest mean correlation coefficients was considered as the match. If the mean correlation coefficient was smaller than 70%, no target was identified. Same principle could be used to incorporate third sound to make the decision making even more robust. The memory requirement for Algorithm 3 was 336.46 MB per 1000 users. The identification and verification times were 648 ± 12 and 3542 ± 131 ms respectively. The accuracy of Algorithm 3 was same as Algorithm 2. It enhanced the speed of identification.

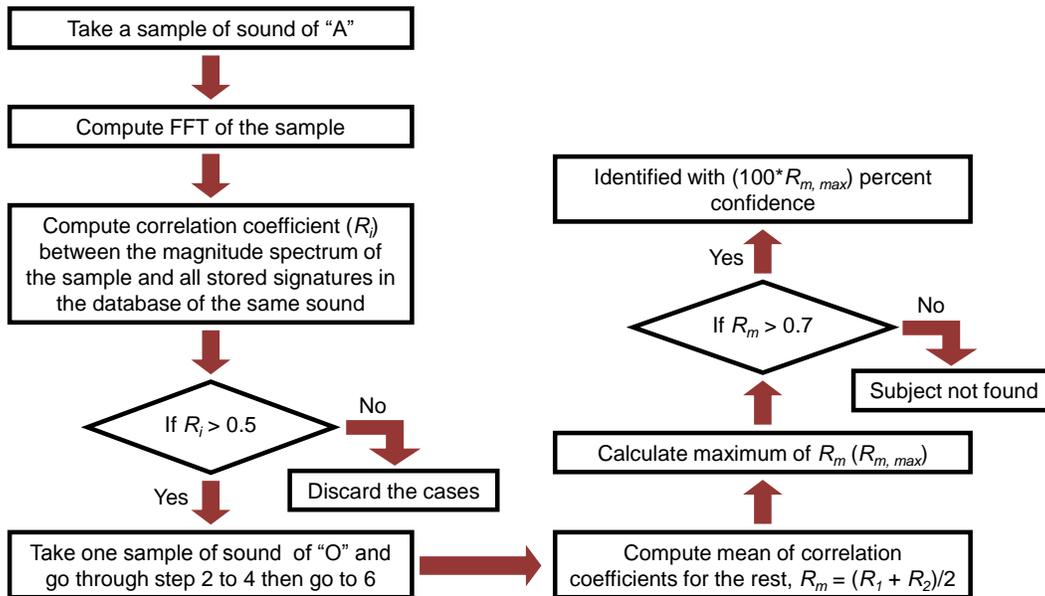


Figure 3-12 Flow chart of Algorithm 3 (identification)

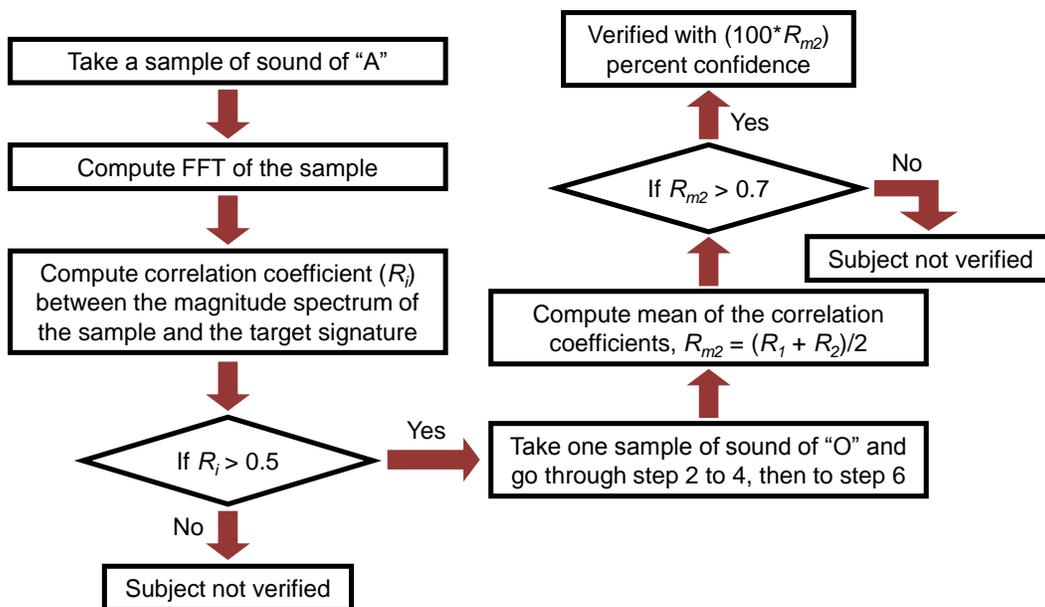


Figure 3-13 Flow chart of Algorithm 3 (verification)

Speaker Recognition using Split Spectrum (Algorithm 4)

Spectrum Splitting

Human ear is an excellent system for sensing sound. It can sense a wide range of frequencies. The audible range of frequencies for human is from 20 to 20,000 Hz. Animals have different audible ranges. For human, the sensation caused by frequency is usually referred to as the pitch of a sound. Human auditory systems don't perceive frequency in a linear manner over the whole audible range. Since frequency is not perceived linearly by human brain, a pitch scale called Mel scale was proposed in 1937 by Stevens, Volkman, and Newman that is linear with human perception [63]. An approximate formula widely used for Mel scale is as below

$$m = 1127 \ln \left(1 + \frac{f}{700} \right) \quad (\text{Equation 3-1})$$

Where f is frequency of a sound in Hz and m is pitch in Mels.

Table 3-2 Relation between frequency and Mel number

	Frequency (Hz)	Pitch (Mel)
Band1	0 – 174	0 – 250
Band2	174 – 391	250 – 500
Band3	391 – 662	500 – 750
Band4	662 – 1000	750 – 1000
Band5	1000 – 1422	1000 – 1250
Band6	1422 – 1949	1250 – 1500
Band7	1949 – 2607	1500 – 1750
Band8	2607 – 3429	1750 – 2000
Band9	3429 – 4454	2000 – 2250
Band10	4454 – 5734	2250 – 2500
Band11	5734 – 7332	2500 – 2750
Band12	7332 – 9327	2750 – 3000

It is an empirical scale of pitches judged by listeners to be equal in distance from one another. A 1000 Hz pure tone with a sound pressure 40 dB above the threshold of a listener was assigned as 1000 Mels. The pitch was almost linearly perceived in the range of 0 to 1000 Hz, and then it became logarithmic above 1000 Hz.

Since the sampling rate was chosen 22050 Hz, i.e. it sampled signals reliably below 11025 Hz according to Nyquist criterion. From the plot in Figure 3-14, it is evident that very little information is present on the magnitude spectrum of Sound "A" above 5000 Hz. Considering 0 to 10 KHz (approx.) as the total spectrum, it is divided into 12 equal bands each with 250 Mel width as shown in Table 3-2. Out of these 12 bands, first 9 bands contain most part of the information of the signal.

Determination of Significant Bands

Given the nonlinearity of the auditory system, it is still capable of separating the speaker and the speech without rigorous effort. This observation was incorporated into the recognition system. Rather than matching the entire magnitude spectrum, it was split according to Mel scale and then these partial spectra were used to find the identity of a speaker.

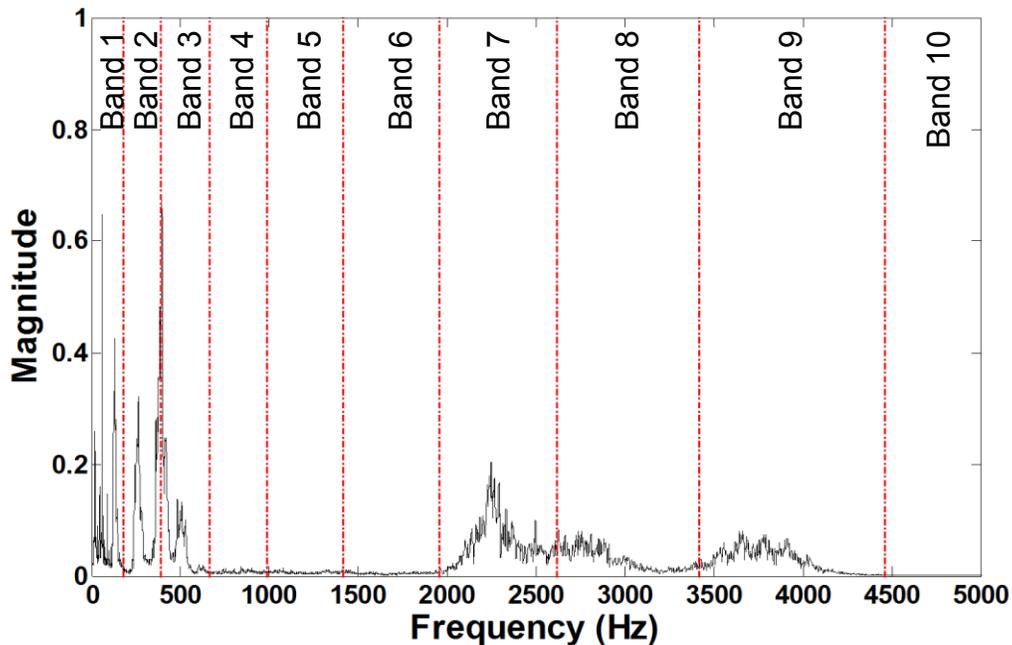


Figure 3-14 Spectrum splitting of a Sound ("A") with an equal Mel interval of 250 Mel

Mel scale splitting of the magnitude spectrum of Sound "A" is shown in Figure 3-14. Each plot of Figure 3-15 is a bar chart with three triplets. Each bar triplet corresponds to correlation coefficient between the samples of a particular subject with three stored signatures. For example, the first, second, and the third bar of the leftmost triplet of Figure 3-15a represents correlation coefficients between the sample of Subject 1 with signatures of Subject 1, 2, and 3, respectively (from the left). Among these 12 frequency bands, each band does not carry equal amount of information as can be seen from Figure 3-14. The information is distributed in a pattern for a particular sound. Some of the bands contain more information than others. From Figure 3-14, it is easily visible that Band 1, 2, 3, 7, 8, 9 carry more information than the rest of the bands. Significance of a band is defined as the information (visible features that are not noise) content available in that band. Given the definition, Band 1 and 3 were more significant than 8

and 9. To find how much information about a speaker was encrypted in a band, correlation coefficients was computed for the samples from three subjects with their stored signatures. Figure 3-15 and Figure 3-16 shows the comparison of correlation coefficients for all subjects. Correlation coefficient and distance are the two important parameters to determine the significance of a band in speaker recognition. Ideally any sample should produce highest matching (correlation coefficient) with its own signature and should not produce higher matching with others when entire magnitude spectrum is used to compute the correlation coefficient. Practically it's a less probable scenario due to the many non-idealities such as noise, variation in our voice, physical illness, etc. Most of the times even if a sample produces highest matching with its own signature, it also produces high matching with few others which makes the decision making less robust. It is always better to have highest matching with its' own signature along with high correlation distance (margin of decision making) with others.

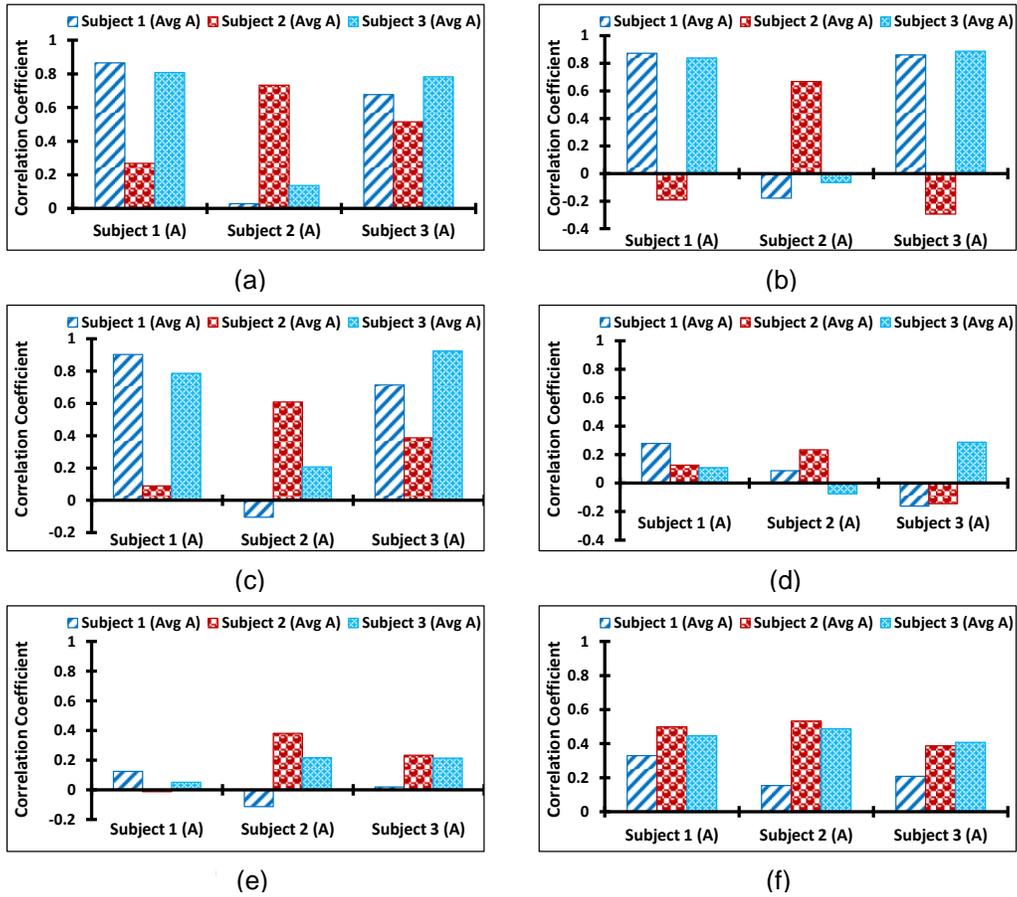


Figure 3-15 Correlation coefficients between a sample of sound “A” of Subject 1 with three stored signatures of three subjects including Subject 1 for different frequency bands; (a) Band 1, (b) Band 2, (c) Band 3, (d) Band 4, (e) Band 5, and (f) Band 6

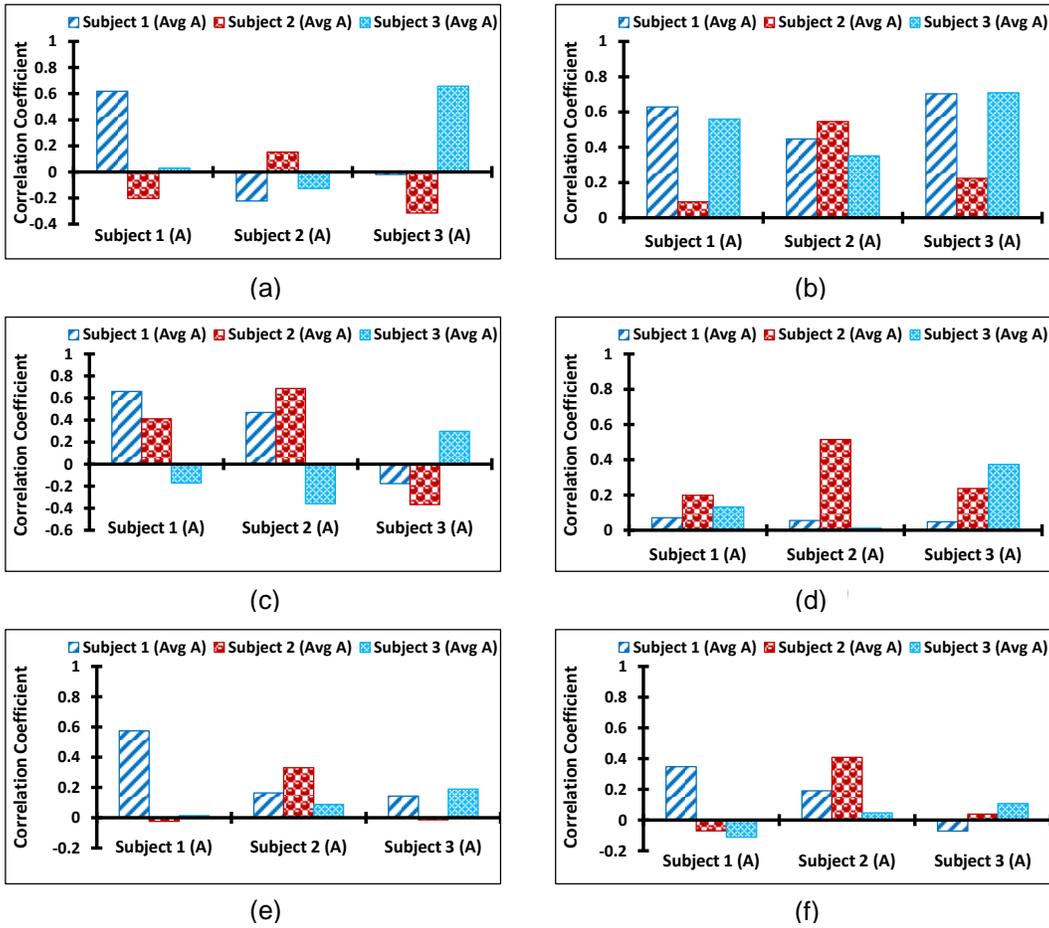


Figure 3-16 Correlation coefficients between a sample of Sound "A" of Subject 1 with three stored signatures of three subjects including Subject 1 for different frequency bands; (a) Band 7, (b) Band 8, (c) Band 9, (d) Band 10, (e) Band 11, and (f) Band 12

Considering matching and contrast, Band 1 to 9 were found to be significant for sound "A" and "O". Similar to Figure 3-15 and Figure 3-16, the correlation coefficients for sound "O" were also computed and the significant bands were determined. A linear weighting was used to assign numeric significance to each band as shown in Table 3-3. Integer weighting was chosen for simplicity. Significance of a band was represented with

a numeric weight. Higher weight meant higher significance. Table 3-3 enlists the significance of these bands for sound “A” and “O” in descending order.

Table 3-3 Significant bands and their weights for Sound “A” and “O”

Sound of “A”		Sound of “O”	
Band	Weight	Band	Weight
3	9	3	9
7	8	1	8
9	7	8	7
1	6	9	6
2	5	7	5
8	4	6	4
4	3	2	3
5	2	4	2
6	1	5	1

Target Identification using Split Spectrum

In most of the significant bands, the target signature produced highest matching with the sample. It failed in few cases though. For example, in leftmost triplet of Figure 3-15f (Band 6), the first bar should be the tallest i.e. sample of Subject 1 should produce highest coefficient with the signature of Subject 1. Instead, it produced highest matching with signature of Subject 2 which was undesirable. As a result, a single band shouldn't be used to detect a target. The results must be combined from separate bands to make the final decision. The natural way of combining results is to calculate average. Since each band has different significance in recognizing a speaker, weighted average was used to combine results. The effect of combining correlation coefficients from multiple bands is shown in Figure 3-17. Most significant band was taken first, then second most significant band was taken as second band, and so on in the Figure 3-17. For both Sounds “A” and “O”, the first few data points were erratic but they get more stable after taking weighted average of multiple bands. The stability is achieved with little expense on the contrast.

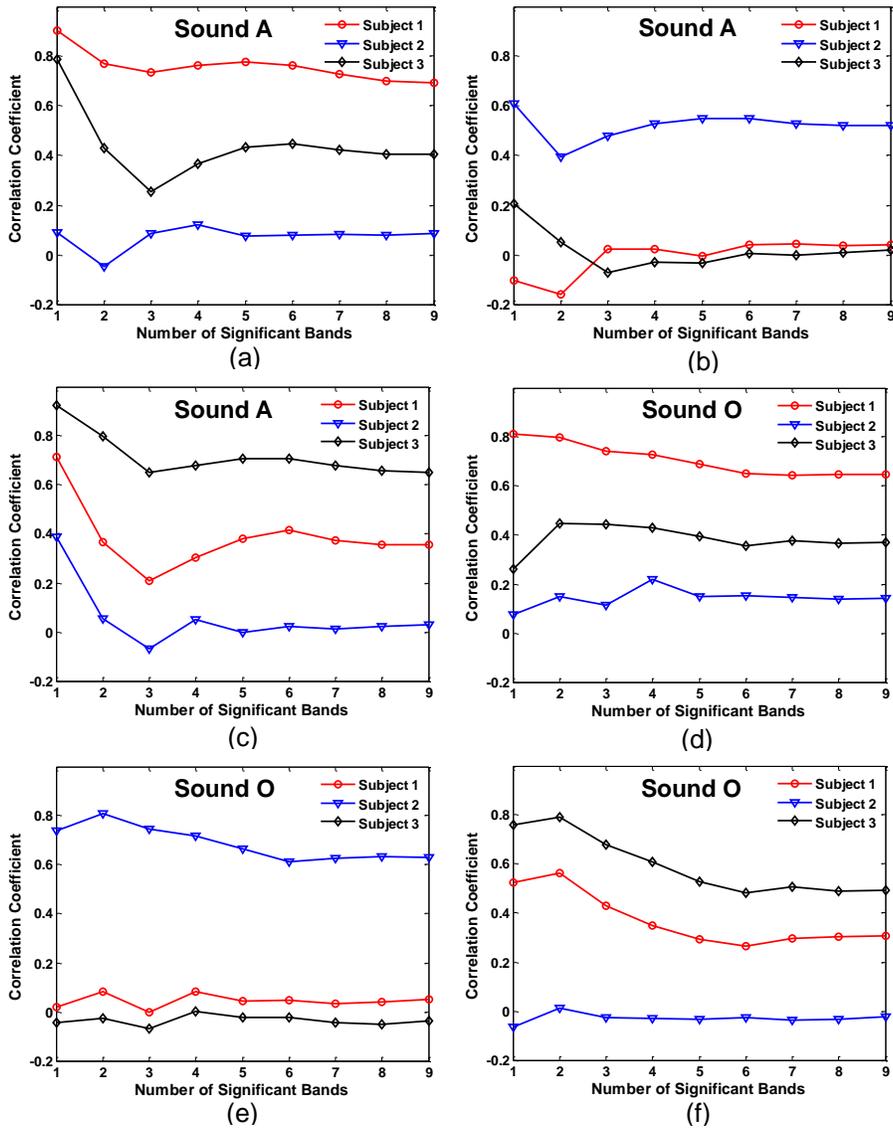


Figure 3-17 Weighted average of correlation coefficients vs number of most significant bands. X-axis represents the number of most significant bands in a descending order to calculate the average and Y-axis represents the weighted average of the correlation coefficients, (a) Sound “A” of Subject 1, (b) Sound “A” of Subject 2, (c) Sound “A” of Subject 3, (d) Sound “O” of Subject 1, (e) Sound “O” of Subject 2, and (f) Sound “O” of Subject 3

During sound production, humans generate specific band of frequencies for a particular sound. The distribution of these bands and frequencies are distinct for different individuals. The whole spectrum was analyzed to find such regions called “significant bands”. For each significant band, the correlation coefficient was computed and then those were used to identify the speaker. The algorithms for identification and verification of Algorithm 4 (split spectrum) are shown in Figure 3-18 and Figure 3-19 respectively. The memory requirement for Algorithm 4 was 168.23 MB per 1000 users. The identification and verification times were 353 ± 1 and 4.938 ± 0.169 ms respectively with an accuracy of 100%.

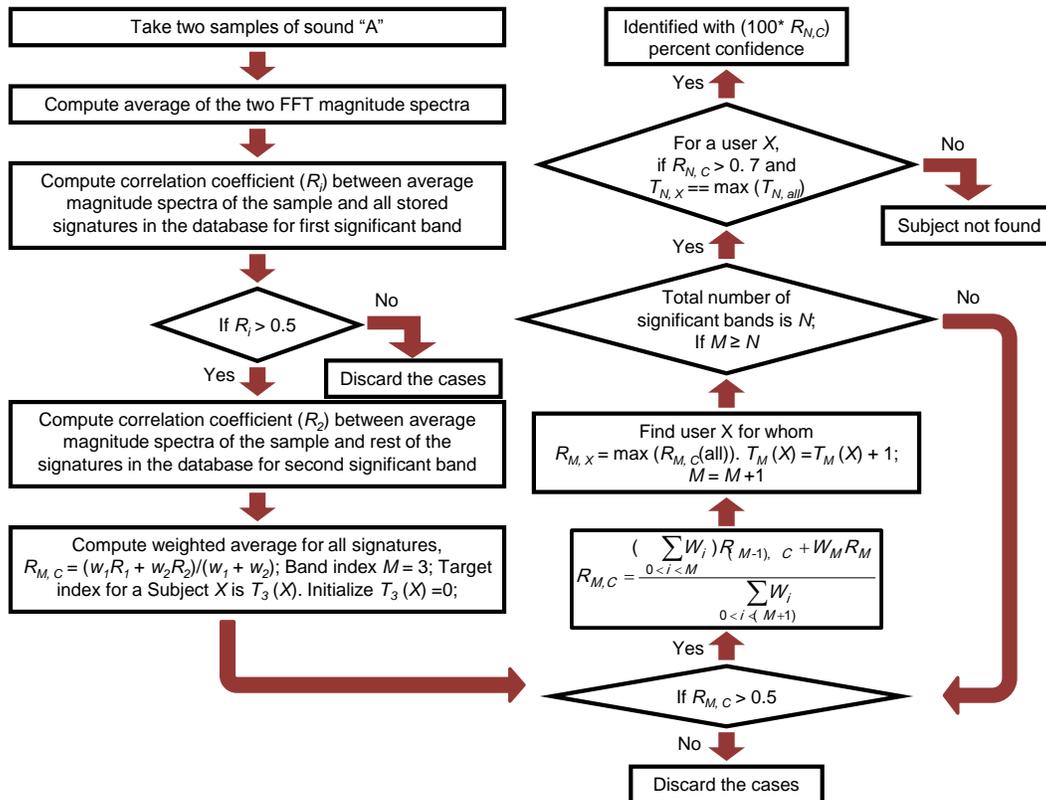


Figure 3-18 Flow chart of the identification using split spectrum (Algorithm 4)

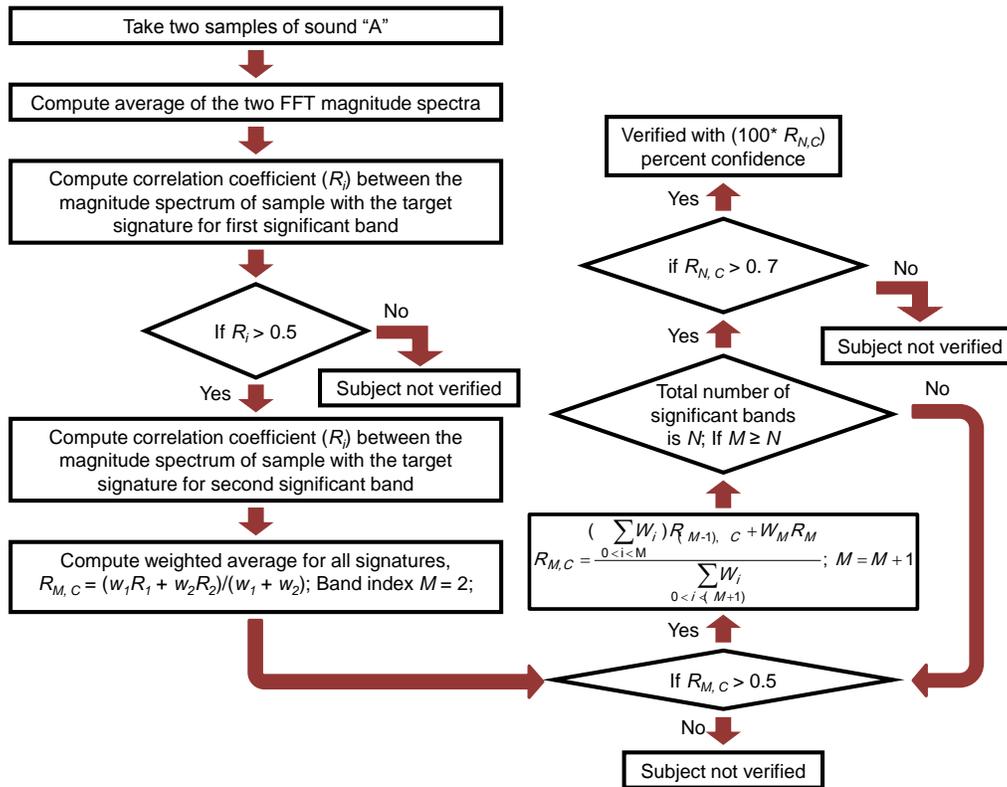


Figure 3-19 Flow chart of the verification process using split spectrum (Algorithm 4)

Comparison

Table 3-4 is a chart of storage and average run time (identification and verification) required for each Algorithm for 1000 users. Algorithm 1 is the simplest approach that is reflected in the memory storage and the average time required to find a hit as listed in Table 3-4. But the accuracy is 93.33%. Algorithm 2 is an improved version of Algorithm 1 in terms of margin of decision making (minimum correlation distance). Higher minimum correlation distance would improve the accuracy when implemented for larger samples. The trade-off is in the database size and average identification time. Since storage has become inexpensive, it is not much of a concern. But, average time required to find a target (average identification time) is very crucial. The accuracy is also very important parameter. The identification time of Algorithm 2 is twice as the

identification time of Algorithm 1. Algorithm 3 keeps the accuracy of Algorithm 2 and reduces the identification time by 42.53%. Identification time for Algorithm 3 is almost equal to that of Algorithm 1. Although the verification times for Algorithm 2 and 3 are almost same, in Algorithm 3, there were ten pairs of sounds as shown in Table 3-5.

Table 3-4 Memory storage, identification and verification time for 1000 users (n = 22)

	Algorithm 1	Algorithm 2	Algorithm 3	Algorithm 4
Memory Storage (MB)	168.23	336.46	336.46	168.23
Identification Time (s)	0.557 ± 0.023	1.126±0.041	0.648±0.012	0.353±0.001
Verification Time (ms)	1.762 ± 0.151	3.532±0.261	3.542±0.131	4.938±0.169
Accuracy	93.33%	96.67%	96.67%	100%

Among those pair of sounds, AO pair (circled) produces the highest matching (84%) with 35% margin of decision making on an average for all subjects. Sound pair “EO” produces 80% matching with the highest margin of 36%. Considering both metrics, overall “AO” would be the best candidate to identify and verify speaker more accurately and faster if Algorithm 3 is implemented.

Table 3-5 Comparison of correlation coefficient and correlation distance

Sound Pair	Average of correlation coefficient for three subjects	Average of correlation distance for three subjects
AE	0.79	0.35
AI	0.81	0.29
AO	0.84	0.35
AU	0.77	0.33
EI	0.77	0.30
EO	0.80	0.36
EU	0.73	0.34
IO	0.82	0.30
IU	0.76	0.28
OU	0.78	0.34

Algorithm 4 required 168.23 MB storage for 1000 users which was same as Algorithm 1. But it reduced the identification time by 36.62 percent compared to Algorithm 1. The enhancement in speed could be achieved further if all the significant bands were not used to calculate the combined average of correlation coefficients. Only first 3 or 4 bands were good enough to isolate the target from a pool of subjects. The identification and verification times were 353 ± 1 and 4.938 ± 0.169 ms respectively. Algorithm 1, 2, and 3 all failed to produce highest matching for one case; Algorithm 4 was able to find the right target every time. As a result, the accuracy reached to 100%. The verification times for all the algorithms were of the order of couple of milliseconds and by definition, these were not dependent on the number of users. As a result, identification time was the most important metric to measure the pace of a system.

Peaks of FFT Magnitude Spectrum Analysis

When computed the FFT magnitude spectra of each sound, it showed many peaks. Each sound was recorded six times. Each subject showed a distinct set of consistent peaks. Consistent peaks were the ones that appear in all samples for a particular sound. The number of consistent peaks (NoCPs) are different for different subjects for a particular sound. Each peak has three distinct features- peak location, peak width, and normalized peak amplitude. Number inside red balloons on all figures (Figure 3-20, Figure 3-21, and Figure 3-22) are representing NoCPs. Location, width, and amplitude of the peaks were distinct for different subjects. They can provide more metrics to make the voice based speaker identification even more robust. Figure 3-20, Figure 3-21, and Figure 3-22 show the distributions of peak locations, peak widths, and normalized peak amplitudes, respectively for 14 subjects. The more the consistent peaks are found in a speaker's sound, the easier it gets to identify that speaker using various

features of the peaks. If a particular feature doesn't appear in all samples of a particular sound (e.g. Sound of "A"), the same feature can be taken from the second sound (e.g. Sound of "O") to improve the recognition.

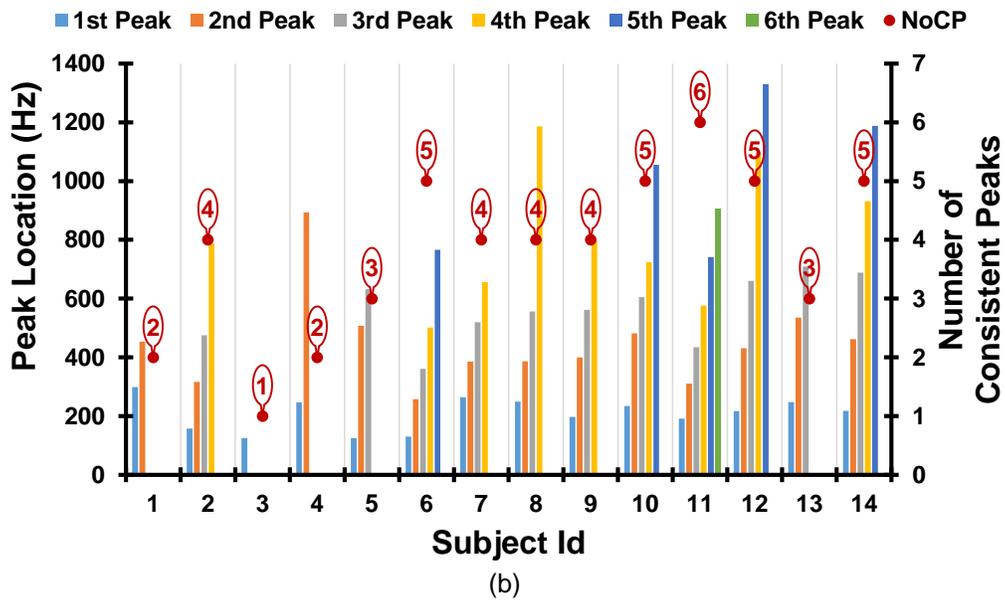
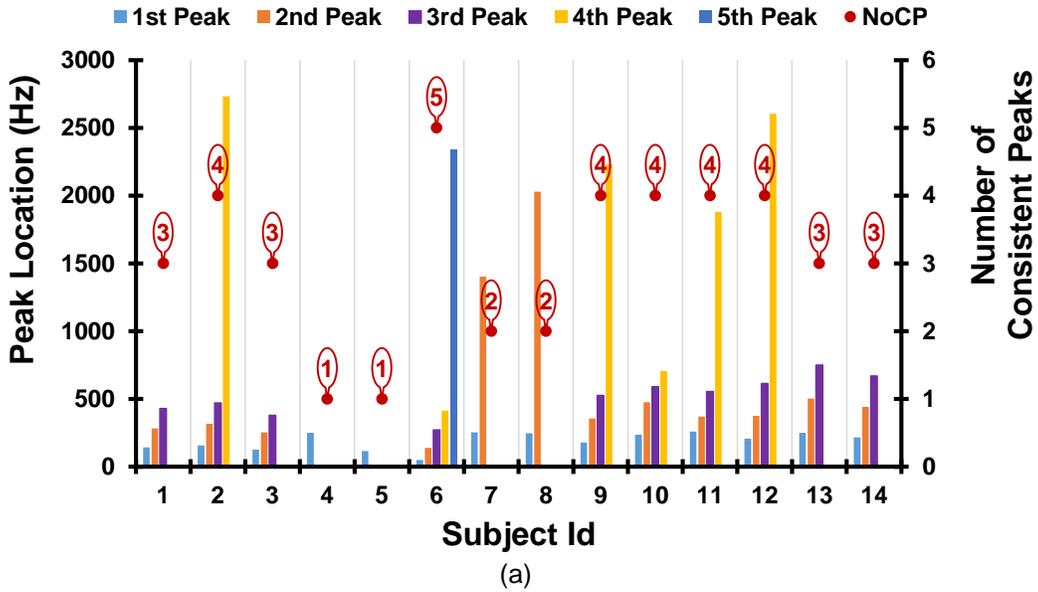


Figure 3-20 Location of peaks for (a) Sound of "A", (b) Sound of "O"

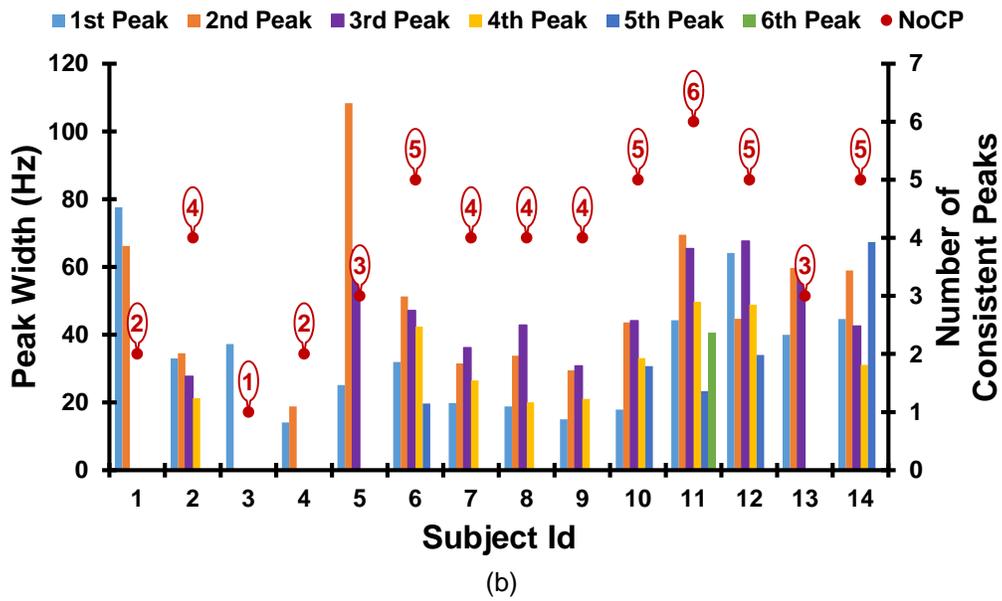
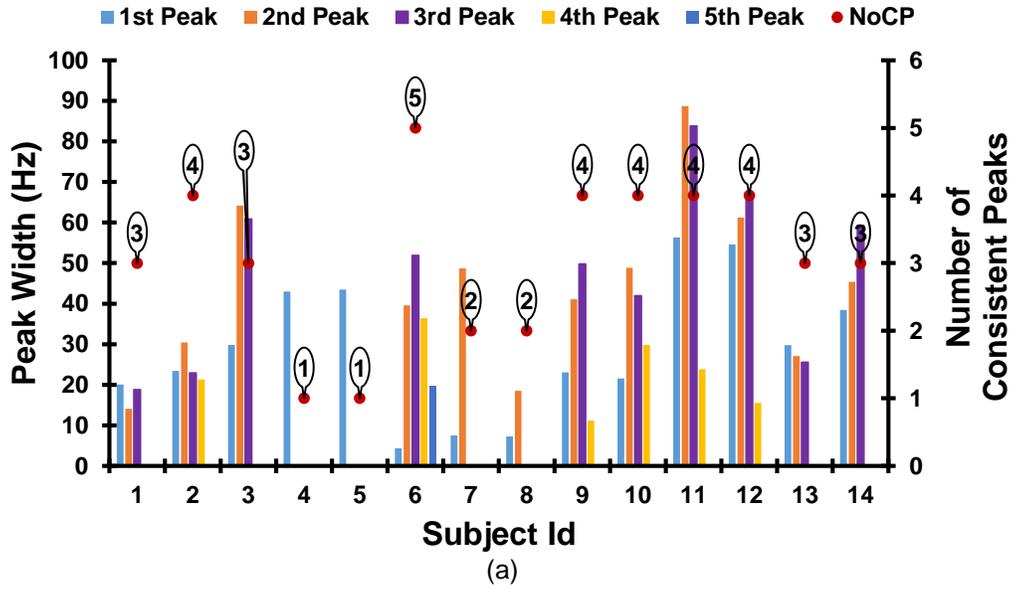


Figure 3-21 Peak widths of (a) Sound of "A", (b) Sound of "O"

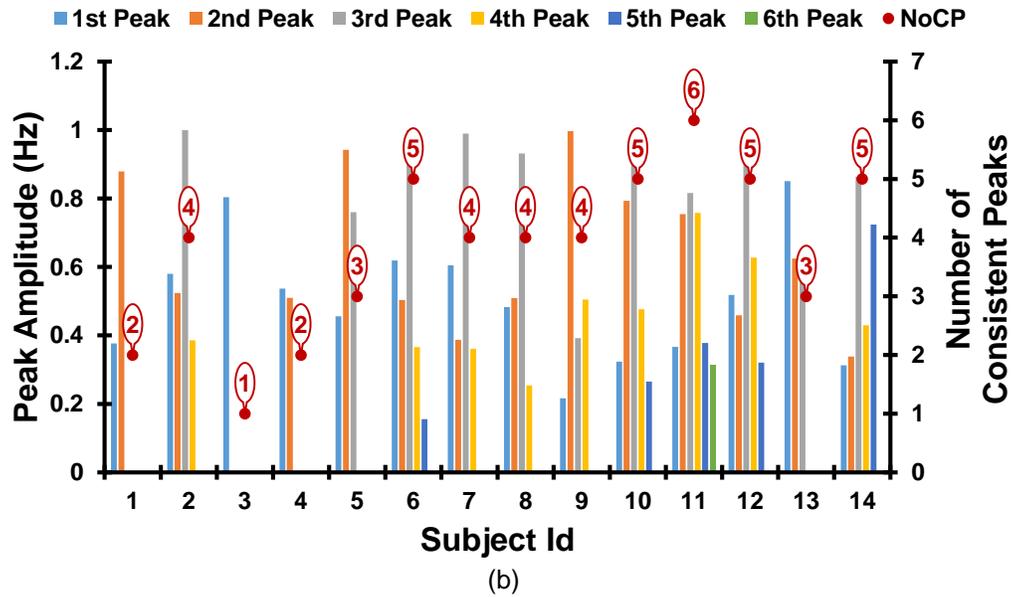
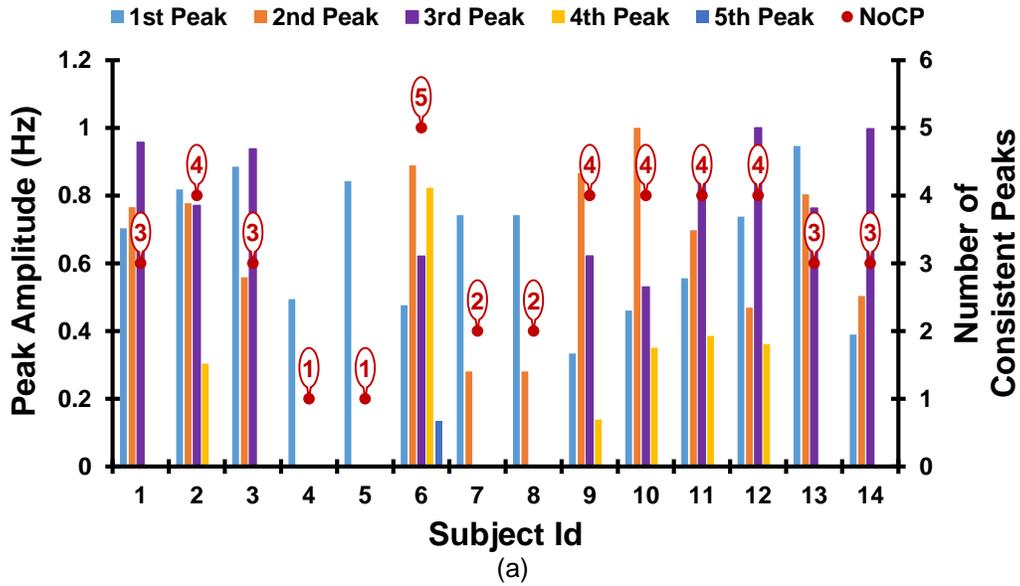


Figure 3-22 Normalized peak amplitudes of (a) Sound of “A”, (b) Sound of “O”

Discussion

Voice based authentication system may replace the traditional passcode based authentication system due to its convenience. The technology has already been used in many areas but not widely accepted yet due to insufficient robustness. With the continuous improvement, it wouldn't be too far in the future when voice based authentication would be integrated in every device we use. Three distinct features of FFT peaks can also be used to distinguish speakers along with the spectrum to make the recognition process more robust. It will be a two way authentication system. One feature will complement others. The computational complexity ($O(n)$) of all algorithms developed here is linear with the number of user, i.e. $O(n) = n$. Since complexity doesn't grow exponentially, the method can be applied to large database.

Chapter 4

Children Speech Recognition and Personal Language Trainer

Introduction

Meaningful speech carries speaker's identity along with the message. The message depends on the encryption method i.e. language. Exact same sound can mean two totally different things in two languages. It has already been discussed how to extract speaker's identity from human voice in Chapter 3. This section discusses how to extract the message content from human speech toward implementing new applications. In the first part, how speech recognition can play a crucial role helping children with speech impediment will be discussed. Second part focuses on an application that helps learner of a new language to master pronunciation accurately.

Speech recognition has been used in many areas. But the application of this technology in some areas is still at its infancy. One such area is to help language therapist who evaluates the sound production of children with speech disorder. Children learn language by hearing from a person and trying to mimic it. Currently, therapists listen and record all the words manually. They keep track of each word pronounced by a kid and evaluate the performance of each word production. There's no automated system to do the job. The process is very tedious, laborious, and prone to human error.

An android application has been developed that uses standard google database and speech recognition technology to listen each kid and keep record of all words. The goal of the application is to provide an automated system that monitors the progress of pronunciation skills of children with speech disorder. It stores words in a database. Words produced by children was analyzed to compute the accuracy of the pronunciation in real-time. The application has the potential to investigate the source location of the speech

disorder. The word delivery by each child can be recorded and quantified to measure progress with minimal human interaction. The key features of the application developed are as follows:

- i) Automation
- ii) Track individual word and overall performance of each user
- iii) Cloud based accessibility
- iv) Personalized voice therapist
- v) Remote supervision
- vi) Real-time feedback
- vii) Complement existing applications

Results and Discussion

The application was written on an open access platform developed by MIT for Android OS. The name of the platform is called MIT App Inventor. The application can be run on any android device. The prototype of the application has the minimal set of functionalities. At this stage, the application has the capability to do voice to text conversion on individual word. There is provision to manually correct the word before storing in database. All practice words can be saved in the local memory of a device. Words are indexed while saving in the database. It allows sorting and easy indexed retrieval of words later. It has the functionality to compute resemblance of a transcribed word with all the words stored in the database. By analyzing mismatch between the target word and the word produced by a kid, important information regarding the language therapy can be determined. Every sound is produced by the incorporation of a different set places in our sound producing apparatus (vocal cords, tongue, teeth, lips, mouth cavity). If it's found that a kid has difficulty pronouncing a particular type of sound,

then it would be possible to work on that organ to improve the sound production. The therapist will have a definitive part to work on.

Trainer engages in normal conversation with a child and the application records it upon tapping the “Record” button on the app. Once the recording is done, the word appears in the top blank box as shown in Figure 4-1a. The word can be modified if not recognized properly by typing. Figure 4-1b shows the pop-up window when invoked the “Record” button. Pressing “Submit” button stores the recorded word in the database as shown in Figure 4-1c. While saving the word in the database, it will appear at the center of the screen momentarily to prompt the user that it’s saved properly. When a user forgets what word was saved in database last time, “Last Item” button retrieves the word from the database. The process is explained in Figure 4-1d.

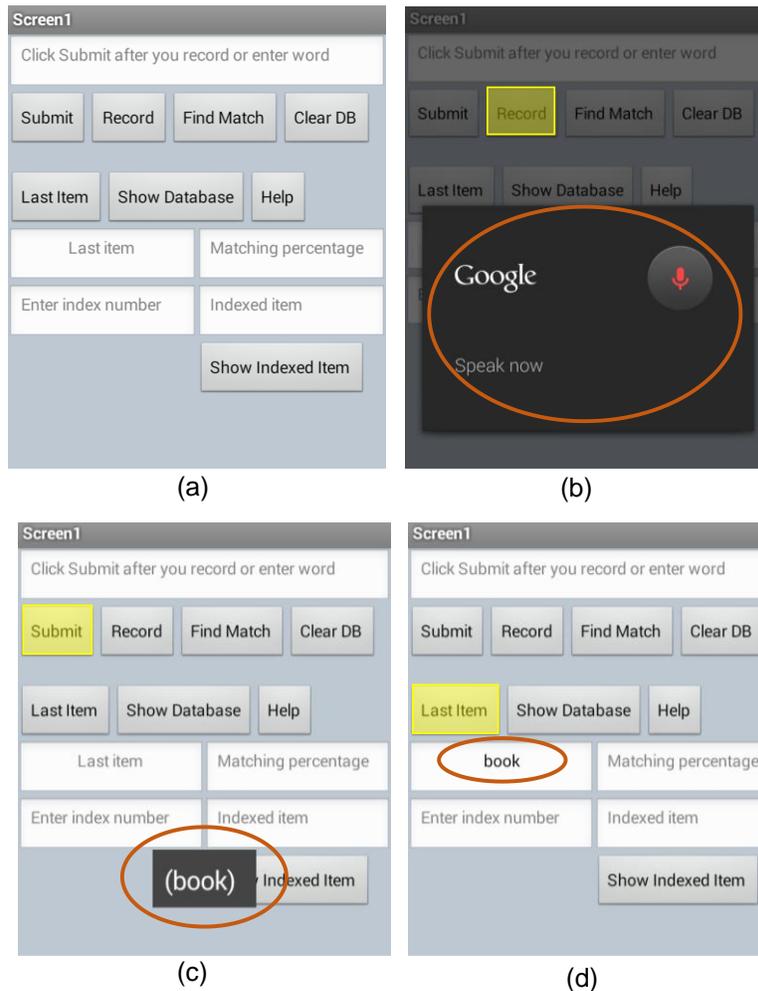


Figure 4-1 (a) Screenshots of home screen, (b) when “Record” is pressed, a small window prompts to speak (circle). Once speaking is finished, it detects the word and shows in the textbox above, (c) Screenshots of home screen while saving a word. Once the word is detected, press “Submit” to save the word in database. The left figure shows a word (book) while saving in the database, (d) when “Last Item” is pressed, it shows what item was stored last in the textbox right below of that button

By pressing “Show Database” button, a user can access all the words stored. It takes the user to another screen (right side) as shown in Figure 4-2a. Figure 4-2b shows how to retrieve an indexed item from the database.

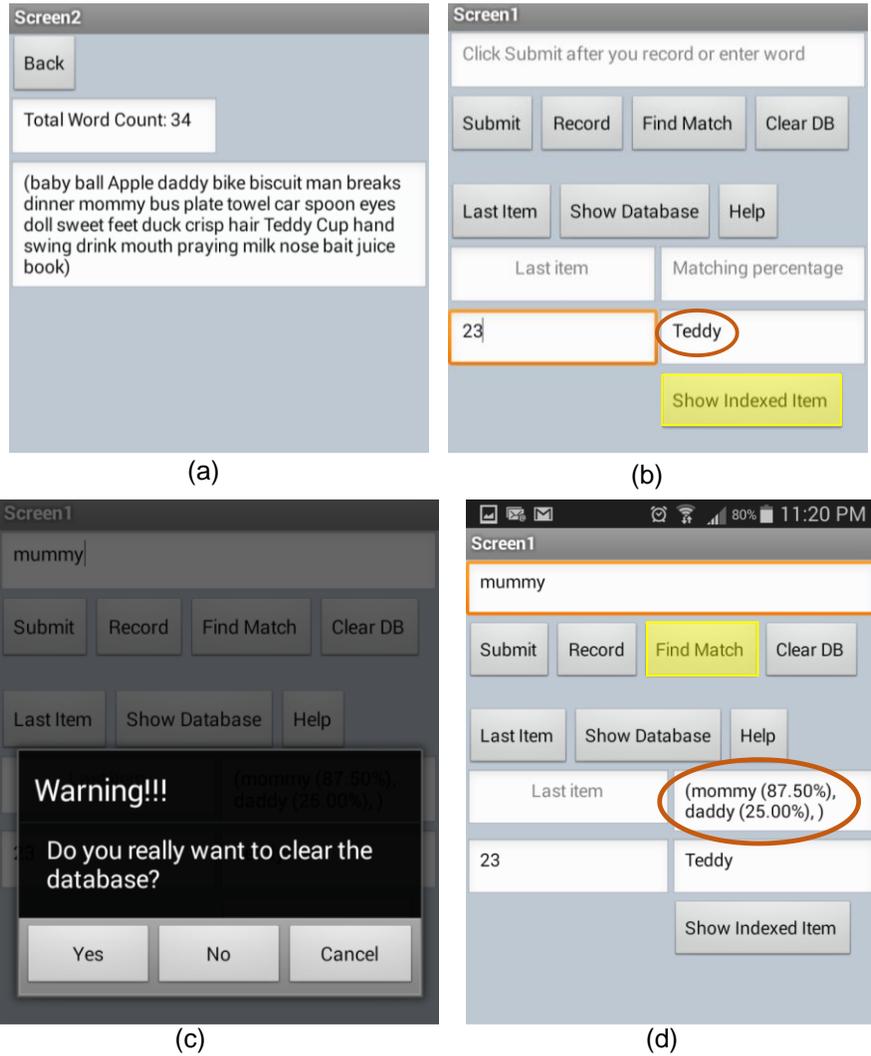


Figure 4-2 (a) Screenshots of the database showing all the words, (b) An index number is entered, the textbox on the right shows the indexed item, (c) Screenshots of a warning message when pressed “Clear DB”, (d) Explaining quantification of how accurately a word was pronounced

When an index number is entered in the text box labeled “Enter Index Number” and “Show Indexed Item” is pressed, it pulls up the item associated with that index number from the database and shows it in a text box right beside the index number. In order to clear the database, “Clear DB” button is used. It will erase everything from the database. Since this button can wipe out all the hard works in a click, a warning message is prompted to the user to confirm the deletion process. It will provide an extra safety layer from losing all data. Figure 4-2c demonstrates this process. A word can be searched throughout the database by entering it in the top text box and pressing “Find Match” button as shown in Figure 4-2d. It calculates the percent match with each word in the database. Once the percentages are computed, it sorts them in a descending order and picks the first two words that produce higher match. A kid was asked to pronounce “mommy” and he produced the word “mummy” which matches 87.5% with a word “mommy” and 25% with the word “daddy” stored in the database. These two words show highest matching with the recorded word. One of the algorithms used normal percentage matching which assigns equal weight on each letter of a word. There is another algorithm that assigns more weight on consonants (1.5) and less weight on vowels (1) to find the matching. It is due to the fact that we hear consonant sounds with more certainty than vowels. The last algorithm assigns highest weight (2) on the first and last letter and then consider nonuniform weighting (consonants 1.5 and vowels 1) for all the other letters in the middle. The justification for this weighted average is we don’t listen all the parts of a word with equal importance. We put more attentions to the terminals sound than to the intermediate ones.

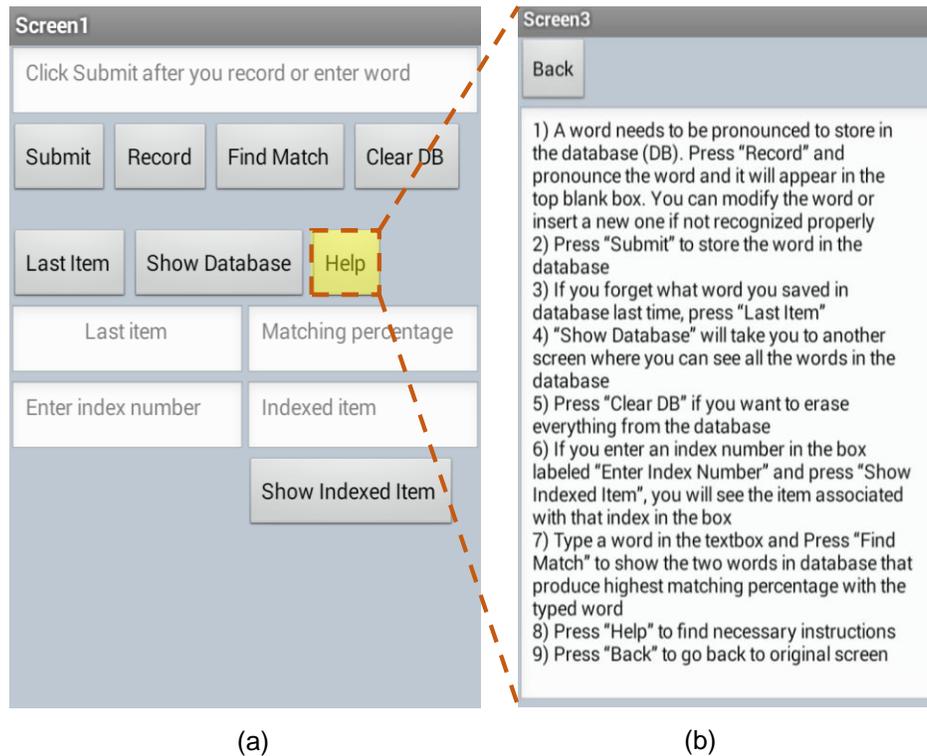


Figure 4-3 (a) Screenshot of home screen with highlighted "Help" button, (b) When the "Help" button is pressed, it prompts a new screen that describes the functionality of each button in the application

The "Help" button is to guide the user to the functionalities of each button in details. It tabulates all the necessary instructions for running the application as shown in Figure 4-3. There are three separate screens used for the application. Each new screen (not home screen) has a "Back" button at the very top of it which will allow the user to back to the home screen (screen 1).

Visual Accent Trainer

What is it?

This is an application package developed on MATLAB graphical user interface (GUI) to help people learn new language in a more engaging and efficient ways. It prompts a user to pronounce a sound in a given language, record the user's input, and then provide a real-time visual feedback on a graphical interface to show how close the attempt was compared to a 'native speaker'. The feedback is quantitative and visual. As a result, a non-native speaker can evaluate his/her performance of the sound production accurately, even if he/she cannot hear the subtle differences in the sound. The speaker can attempt to copy each word multiple times trying his/her sound production closer to the target.

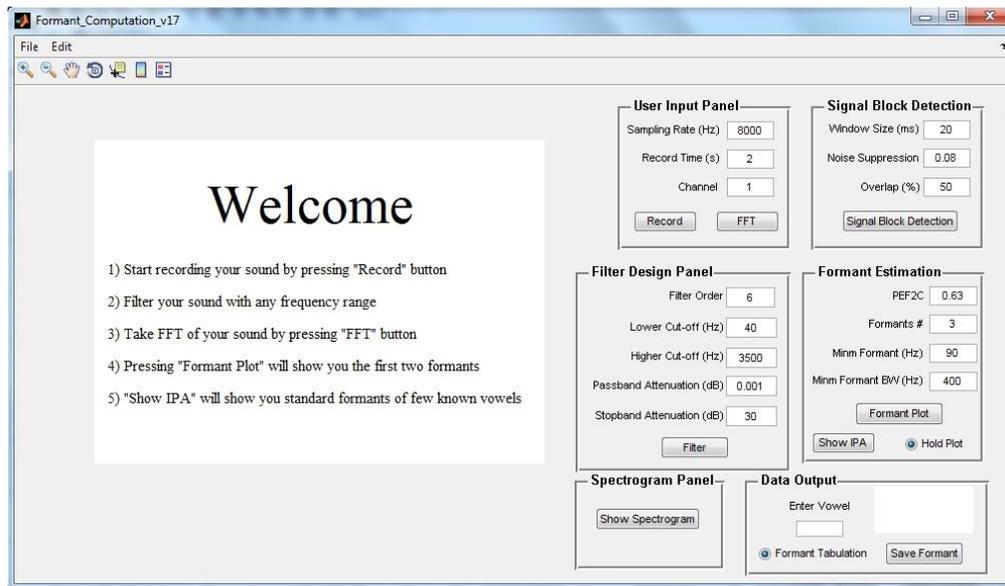
Formant Frequency

According to Benade (1976), formants can be defined as the peaks of the spectrum envelope of a sound [64]. Acoustical Society of America defines formant as a range of frequencies in which there is an absolute or relative maximum in the sound spectrum. Formants are essentially the resonance frequencies of a sound. In the process of sound production, a puff of air is pushed upward through the laryngeal tube. The vocal cords chop the air with a certain frequency and the vibrations produced act as the fundamental frequency of that sound. The sound is later shaped by mouth cavity, tongue, teeth, and lips. The fundamental frequency produces multiple resonances in the mouth cavity depending on the position of tongue, teeth, and lips.

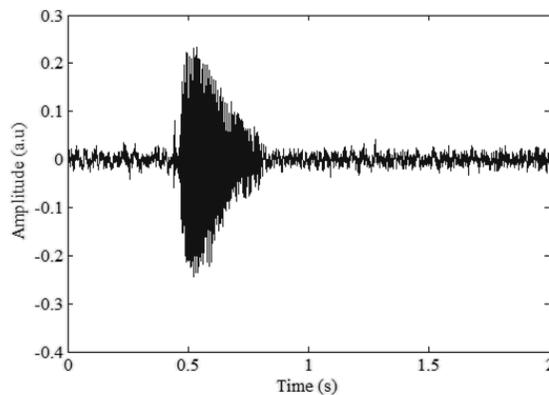
Methods

User Guide for the Application

- i) Once the application is run, a home screen pops up with a set of instructions on the front panel as shown in Figure 4-4a. Hovering mouse pointer to any button shows a little box with more information about that button.



(a)



(b)

Figure 4-4 (a) Screenshot of home screen of the MATLAB graphical user interface (GUI) application, (b) Time domain signal of the sound of English vowel “A”

- ii) All the values displayed in the text boxes are default values. Any of them can be tweaked to get desired results.
- iii) Pressing the “Record” button starts recording. Please wait till you hear a dot sound (make sure the speaker and microphone are turned on).
- iv) After recording the sound, the display window looks like Figure 4-4b.

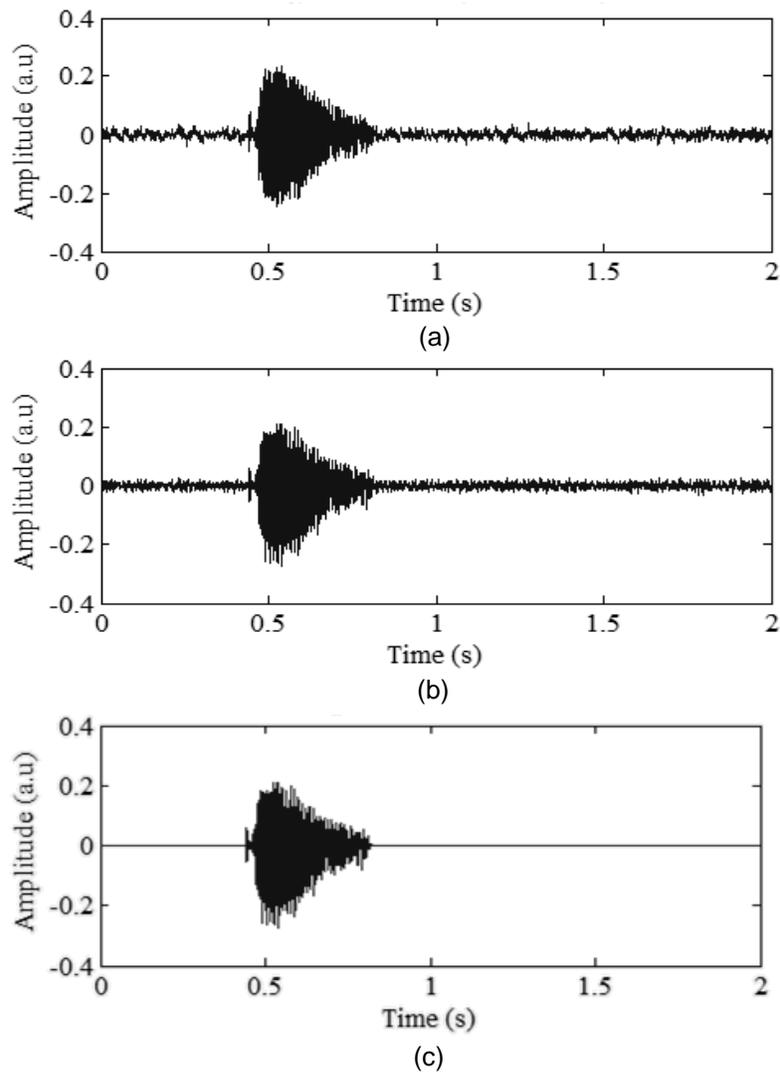


Figure 4-5 (a) Time domain signal of sound of “A”, (b) Filtered version of the same signal ($f_{CL} = 40$ Hz and $f_{CH} = 3500$ Hz), and (c) Automatically detected signal block

v) Pressing the “Filter” button removes noise from the signal. A six order elliptic band-pass filter was used here. It filters out any unwanted signal that falls outside of the band 40 – 3500 Hz. The cut-off frequencies were so chosen due to the fact that most frequency components of the vowels lie well within this range. Attenuation for pass and stop bands were 0.001 and 30 dB respectively. Figure 4-5a, Figure 4-5b, and Figure 4-5c show a signal, its filtered version, and detected signal block of the filtered version, respectively.

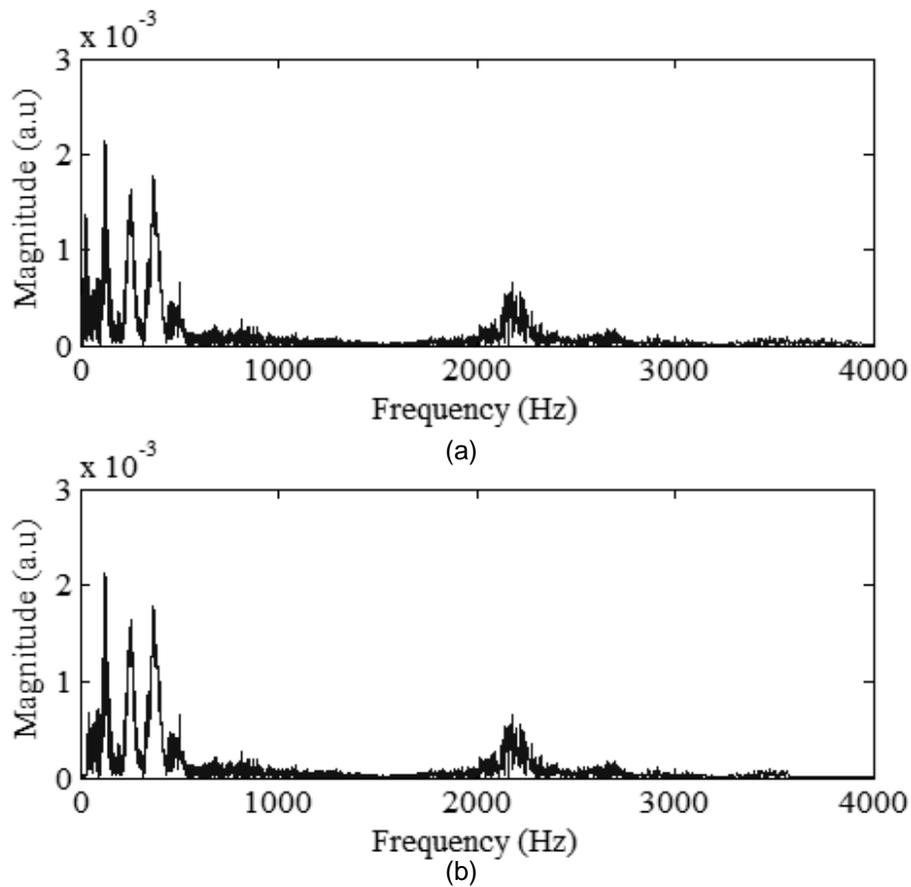


Figure 4-6 (a) FFT magnitude spectrum of sound of “A”, (b) FFT magnitude spectrum of the filtered version of the same signal

- vi) "Signal Block Detection" button is to detect exactly when the signal block appears in the time series as shown in Figure 4-5c.
- vii) Pressing "FFT" button performs fast Fourier transform on the signal. Figure 4-6a and Figure 4-6b show the magnitude spectrum of the original signal and that of the filtered signal, respectively.

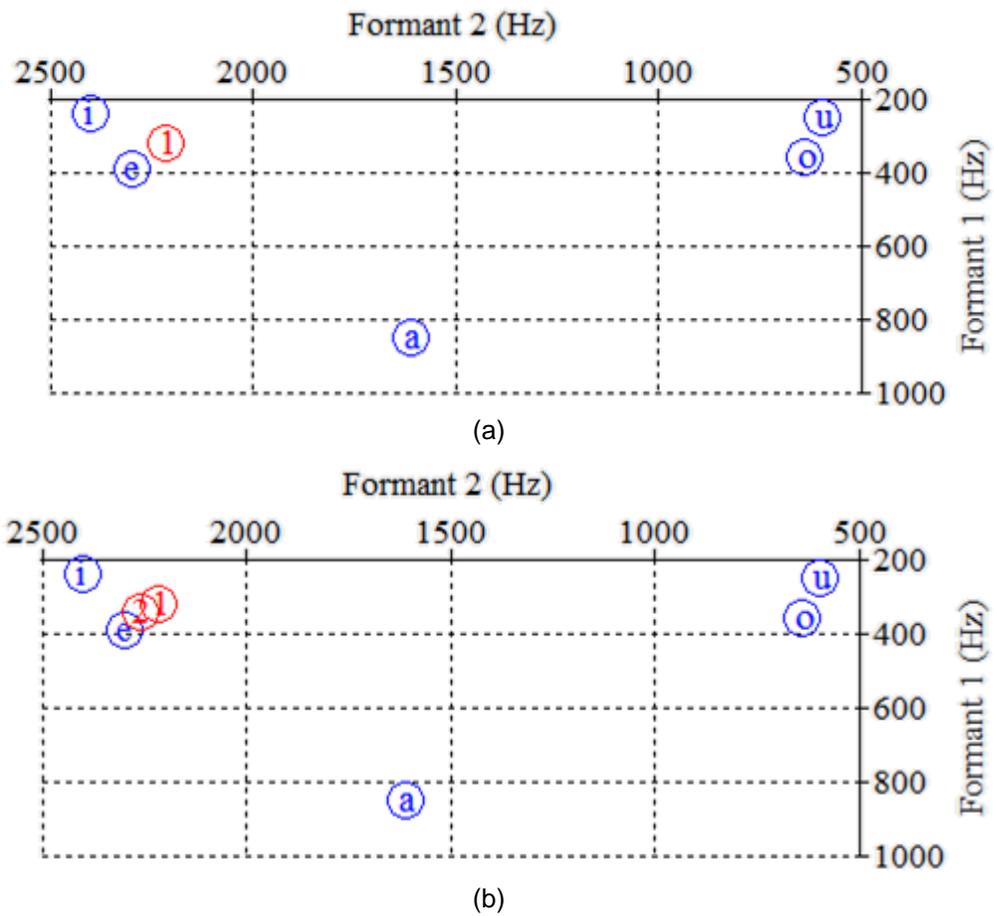


Figure 4-7 (a) Parameterization of Sound "e" trial one, (b) trial two

- viii) Name of the vowel is required to put in the text box (left) on the "Data Output" panel as shown in Figure 4-4a. Once a name is entered in the text box mentioned above,

pressing “Formant Plot” will prompt to save the formant frequencies in a excel file.

This name is used to index the data.

- ix) Once pressed the “Formant Plot” button, a red circle representing the formant frequencies (first two formants) will be appeared on a 2D plot as shown in Figure 4-7a. There's a number inside the red circle which is the sequence number of the recording. There are five other data points on the plot (blue circles with a letter inside) which are the standard locations of the five vowels [a] as in bat, [e] as in bed, [i] as in tee, [u] as in coo, O[o] as in code. The input vowel is the red circled one. First three formant frequencies will be displayed in a text box (right) on the “Data Output” panel which is located on the bottom right corner of the screen. A learner can try multiple times to bring his/her sound production closer to the targets as shown in Figure 4-7b.
- x) If a user unselects the radio button “Hold Plot”, the previous data point will be cleared from the plot, showing only the current data point. Keeping this button on will allow you to hold the consecutive formant points on the plot to compare among them as shown in Figure 4-7b.
- xi) “Show Spectrogram” button allows a user to observe the spectrogram of the sound as shown in Figure 4-8. Due to the computation loading, this process usually takes a little longer (around 22 second).

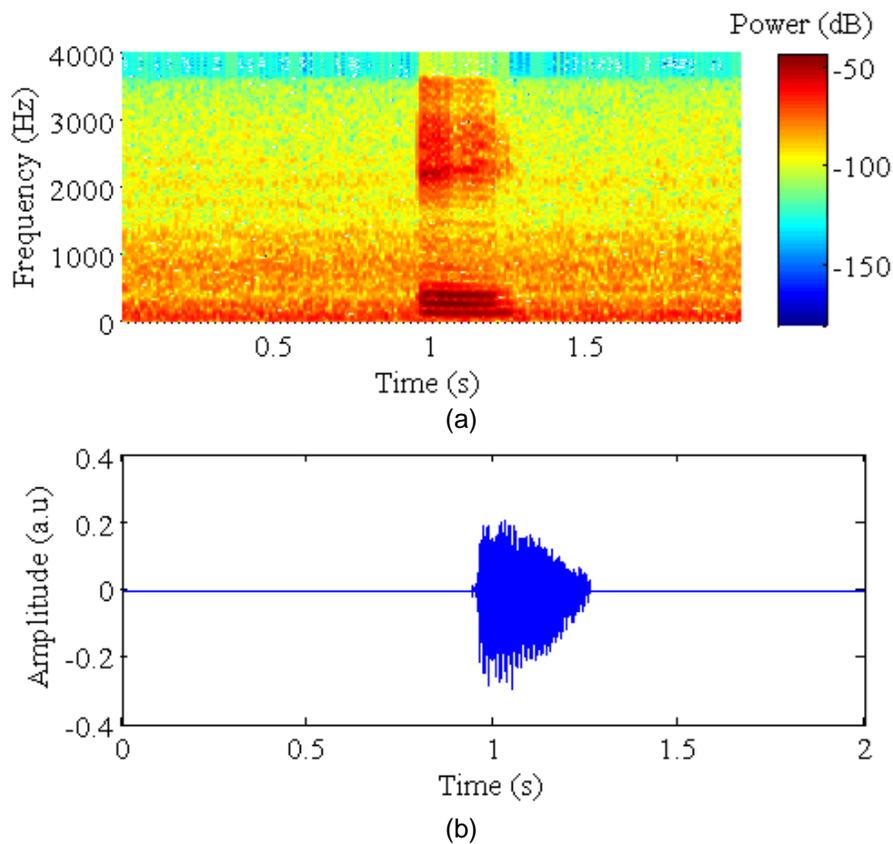


Figure 4-8 (a) Spectrogram of sound of “A”, (b) Detected signal block

- xii) Keeping the “Formant Tabulation” radio button on allows the user to store formant frequencies in a matrix which can be saved later by clicking “Save Formant” button (located on the “Data Output” panel). It prompts the user to select a file name and location and save the data in .xlsx format.
- xiii) Any displayed image can be saved by clicking “File” and “Save as”. It allows the user to choose a suitable output format (png, jpeg, and emf). There is no functionality of the button “Show IPA” and “Edit” at this moment. They will have functionalities in future version.

Results and Discussion

What Problem it Solves?

Learning a foreign language is difficult and rewarding at the same time. But there are very few tools on the market to help a speaker to learn new language without accent. Having a trainer is helpful but it doesn't solve all problems. Usually people from different geographical location can't distinguish many subtly different sounds from other languages, particularly vowel sounds since they are the fundamental building blocks of a language. This ability is lost by the age of approximately 18 months in humans if not exposed to non-native sounds. The situation gets even worse when speaker's native language is very different from the one he/she is trying to learn. To learn how to pronounce the sound accurately, the speaker needs to hear the sound. But ironically he can't hear the difference that is why he can't pronounce the sound properly at the first place. Even though the trainer pronounces distinct sounds (for example "live" and "leave"), the speaker doesn't hear any difference. It makes very difficult for a foreign language learner to hear the difference between the target sound and their own production, leading to a foreign accent when speaking the new language. Without a visual feedback, this vicious loop of learning impediment can't be broken. This application is the first one of its kind that provides visual feedback to the speaker in real-time, allowing them to practice until they reach their target.

The software can also be used to the deaf community who wish to use oral language. It provides them with visual feedback that allows them to train their accents in the same way that the hearing population can do theirs. Parents can also utilize this application to train their kids from very early age so that they don't have accent when they grow older.

Comparison with Similar Existing Technologies

Most of the existing language learning applications only plays words in a foreign language and asks the user to mimic it. At best, few of them let the speaker know that the production is not accurate. But none of them provide any quantitative feedback of the sound production. As a result, most of the time the speaker doesn't know whether his production is correct or not. If he is informed that the production is incorrect then there is no quantitative feedback that tells the speaker how off he was from the target or what modification he needs to make to improve his sound production. In a way, this application acts as the natural auditory feedback loop that we use as infants to learn our native language. Not only the application provides a quantitative feedback, it does provide it visually. Moreover, it can recommend the speaker to work on a particular part of his sound producing apparatus since each vowel is produced by the participation from different parts in our mouth and laryngeal tube.

Chapter 5

Smartphone based Blood Pressure and Heart Rate Monitoring System

Introduction

Cardiovascular (CV) disease is the leading cause of death in US for last seven decades. It refers to any complication associated with the heart, the blood vessels (arteries, capillaries, and veins) or both. Although the mortality rate has been reduced significantly during this time due to advances in monitoring devices, better diagnostics, awareness, and better medications, it still remains at the top of the fatal diseases. The CV diseases responsible for the most human fatalities are myocardial infarction (heart attack), hypertensive heart disease (high blood pressure), sudden cardiac death (SCD), angina, coronary artery disease (atherosclerosis), congestive heart failure, cardiomyopathy, cyanotic heart diseases, heart failure, arrhythmia, and endocarditis [65].

Cardiac Arrhythmia

It refers to any heart conditions with an abnormal rate or rhythm of the heartbeat. It can cause the heart to beat too fast, too slow, or with an irregular rhythm. Any problem with the heart's electrical process can cause an arrhythmia. For example, electrical signals travel through the atria in a fast and disorganized way in a common type of arrhythmia called atrial fibrillation. In this process, the atria start quivering instead of contracting. A heartbeat that is faster than normal is called tachycardia and a heartbeat slower than normal is called bradycardia. Many arrhythmias are harmless but some of them may cause cardiac arrest. The heart may not be able to pump enough blood to the body during an arrhythmia. Lack of sufficient blood flow can damage the brain, heart, and other organs.

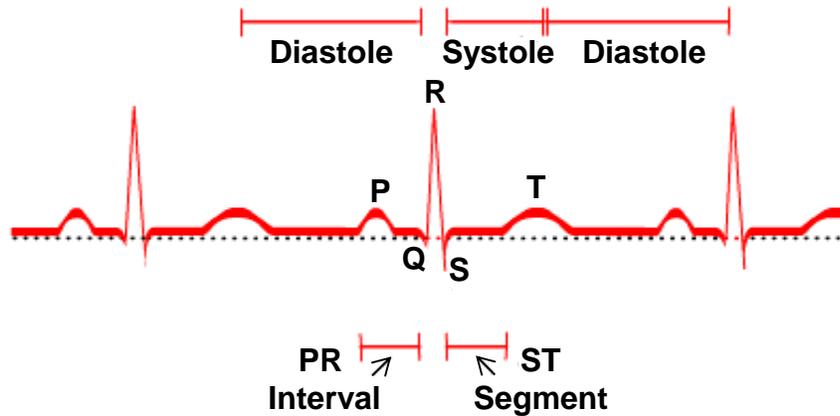


Figure 5-1 A typical ECG signal [66]

A typical ECG tracing of the cardiac cycle (heartbeat) as shown in Figure 5-1 consists of a P wave, a QRS complex, a T wave, and a very small U wave (not visible in Figure 5-1). When the period between Q and T phases of cardiac cycle gets prolonged, it's called long QT syndrome (LQTS). It is a rare inherited heart condition in which delayed repolarization of the heart following a heartbeat increases the risk of a form of irregular heartbeat that originates from the ventricles known as torsades de pointes (TDP). Mutations in the hERG gene cause LQTS. A blockage of hERG channels by various group of drugs can induce arrhythmia [67].

Hypertension

Hypertension or high blood pressure is a condition when the force of the blood against your artery walls is higher than normal. Normal blood pressure (BP) varies depending on the age, gender, mood, heart condition, etc. BP is determined by the amount of blood heart pumps and the amount of resistance to blood flow posed by the arteries. With the increase pumping and the narrower arteries, the blood pressure would be higher. High blood pressure (hypertension) may show no symptom for years.

Sustained high blood pressure increases the risk of serious health problems, including heart attack and stroke. Hypertension can be detected easily if monitored carefully and frequently. Preventive measures can be taken by changing dietary habits, exercising, and changing other life styles (smoke and alcohol). Table 5-1 classifies normal and high blood pressures for adults. Blood pressure and heartbeat measurements on a regular basis play a vital role in the diagnosis of many heart conditions. Cardiac arrhythmia and hypertension are two heart diseases that can directly be diagnosed with heartbeat and BP monitoring. An ideal monitoring system should be portable, easy-to-use, and affordable. Due to prevalence of smartphones now a day, any smartphone based BP and HR measuring system can be used as readily available monitoring devices. Frequent monitoring of heartbeat and BP provides a much more accurate information about health conditions than infrequent readings from doctor's office. But, the current methods of using cumbersome inflatable cuff and electrocardiogram (ECG) do not allow this due to their lack of portability.

Table 5-1 Classification of blood pressure for adults [68]

BP Classification	Systolic BP (mm Hg)	Diastolic BP (mm Hg)
Normal	<120	and <80
Prehypertension	120–139	or 80–89
Stage 1 hypertension	140–159	or 90–99
Stage 2 hypertension	≥160	or ≥100

Turning a smartphone into a health monitoring device has many inherent advantages. Data can be stored in the device as well in the cloud storage where a doctor/clinician can review it. Long term trend can be found without the help of a clinic. Physicians can have much more comprehensive data than before.

Significant numbers of individuals with hypertension are unaware of their condition just because they don't have access to the monitoring tools frequently. If diagnosed at early stages, fatalities due to hypertension can be significantly reduced. Appropriate measures need to be taken to prevent the development of hypertension and to improve awareness, treatment, and control of hypertension in the communities. With an accessible and affordable BP and heartbeat monitoring tool, it is possible to prevent hypertension and thereby save many lives.

Additionally, any self-blood pressure monitoring system is believed to be able to reduce the white coat hypertension effect, elevation of BP during the examination process due to nervousness and anxiety caused in the clinical settings. It is highly recommended to measure BP at home by almost all hypertension organizations. This would eliminate white coat hypertension effect [69]. The tracking of BP from childhood to adulthood has many important implications on public health. There is a growing consensus among the experts in the field of hypertension studies that it starts early in life [70-72]. Recent studies show a relation between the increasing obesity epidemic and high blood pressure in children [73, 74]. Some studies also show that most of the CV diseases except congenital ones can be prevented through better life style (food habit, exercise, regular monitoring, etc.) [29]. Blood pressure and pulse rate are the two most important biomarkers for most of the cardiovascular diseases [75-78]. A smartphone based system was developed that monitored heart rate and pulse pressure (difference between systolic and diastolic pressure) in a convenient way. Fingertip was placed in front of a smartphone camera to record the video for 20 seconds. Then, from the intensity analysis of each frame, heart rate and pulse pressure were calculated.

Background

Heart is the central organ of human cardiovascular system, located in the thoracic cavity (between two lungs). A dual layered connective tissue sac protects and supports the heart. The inner side of the heart wall is formed with cardiac muscle. Each side of the heart has an atrium that receives blood from veins and a ventricle that pumps blood out to the arteries. There are four chambers in total, left and right atria and left and right ventricles, as shown in Figure 5-2. The superior and inferior vena cava are the two largest veins that deliver oxygen-poor blood from the body to the right atrium and the pulmonary veins deliver the oxygen-rich blood to the left atrium. Atrium and ventricle of each side are separated by atrioventricular (AV) valve that ensures unidirectional blood flow. When the blood pressure is higher than a threshold, the valve opens and automatically shuts off when the pressure falls.

Cardiac Cycle

Heart is a pulsatile periodic pump. It completes a full cycle with a sequence of contraction and relaxation. Each complete cycle is called cardiac cycle or heartbeat as shown in Figure 5-2.

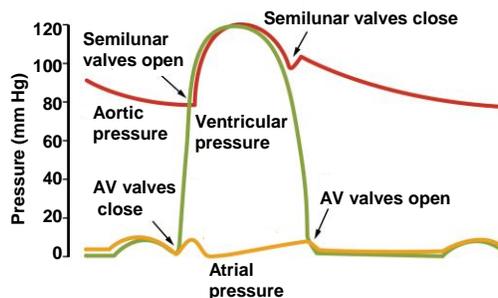


Figure 5-2 Cardiac cycle [79]

A set of specific events occur in each cardiac cycle. The major steps of a complete cardiac cycle are detailed below.

- 1) Both atria and ventricles are relaxed and blood entering into atria forces the AV valves opens. Ventricles starts filling with blood. P wave of the electrocardiogram (ECG) corresponding to the electrical depolarization of the atria causes the contraction of the atrial musculature. As the atria contracts, it fills the ventricles with more blood as the ventricles are still relaxed. Blood does not flow back into the vena cava due to two mechanisms- the inertial effects of the venous return and the milking effect [80].
- 2) This phase of the cardiac cycle coincides with the appearance of the QRS complex of the ECG that represents ventricular depolarization. It triggers the myocytes to contract rapidly with a result of increased intraventricular pressure which in turn closes the AV valves. Pressure keeps building inside the ventricles until the pulmonary and aortic valves open. Ventricular contraction also causes papillary muscles to contract with their attached chordae tendineae that prevent the AV valve leaflets from bulging back into the atria and becoming leaky. The first heart sound “lub” is caused by the closure of the AV valves. This sound is a mixture of two shortly apart sounds due to the fact that mitral valve closes slightly before earlier than tricuspid valve [81].
- 3) When the ventricular pressure exceeds the thresholds of pulmonary and aortic valves, it causes a rapid ejection of blood into the aorta and pulmonary arteries from the ventricles. Outflow velocity, aortic, and pulmonary artery pressures are at their maximum level early in the ejection phase.
- 4) The beginning of ventricular contraction (approximately 200 ms after the QRS complex), ventricular repolarization occurs (T-wave of the ECG) which leads to a

- decline in the rate of ejection. Outward flow continues due to inertial energy of the blood even though ventricular pressure falls slightly below outflow tract pressure. Continued venous return from the lungs and from the systemic circulation cause gradual increase of the left and right atrial pressures respectively [80].
- 5) The aortic and pulmonary valves closes abruptly (aortic precedes pulmonic) at the end of phase 4 when the intraventricular pressures fall sufficiently. It starts beginning of isovolumetric relaxation along with the second heart sound (S2). A small backflow into ventricles generates a characteristic notch (incisura or dicrotic notch) in the aortic and pulmonary artery pressure tracings. The aortic and pulmonary artery pressures rise slightly (dicrotic wave) after the valve closure followed by a slow decline in aortic pressure [80].
 - 6) As the ventricles continue to relax, the intraventricular pressures drops below their respective atrial pressures at some point. It causes the AV valves to open and start ventricular filling. Ongoing ventricular relaxation causes the intraventricular pressure to dip in spite of the inflow of blood from the atria. As soon as the ventricles are completely relaxed, the ventricular pressure starts rising as blood fills in. Ventricular filling doesn't produce any sound normally. But when a third heart sound (S3) is recorded, it may represent tensing of chordae tendineae and AV ring during ventricular relaxation and filling.
 - 7) With the continual filling of ventricles, their compliance reduces and the intraventricular pressures rise. Due to the drop in pressure gradient across the AV valves, the rate of filling in drops. In a healthy resting hearts, the ventricle is filled 90% of its' capacity by the end of this phase. Pulmonary arterial and aortic pressures continue to drop during this phase.

Heart's Electrical System

Each heart beat is initiated with an electrical impulse from a group of cells called sinoatrial node (SA) located in the right atrium. Once the electrical signal is generated, it passes through special pathways in the right and left atria. This causes the both atria to contract and pumps blood into the heart's two lower chambers, the ventricles. In between the atria and the ventricles, there's a group of cells called the atrioventricular (AV) node. The electrical signal then travels down to AV node and slows down a little in AV node that gives the ventricles enough time to finish engorging with blood. The signal then travels through a pathway called the "bundle of His". The bundle of His divides into two branches. The left and the right bundle branches run along the interventricular septum. The left branch further divides into the left anterior and the left posterior fascicles. Purkinje fiber is a thin filament arisen from these bundles and fascicles. It distributes the impulse to the ventricular muscle causing them to contract and pump blood to the lungs and the rest of the body. The ventricles then relax, and it starts the whole process again in the SA node. The electrical signal travels through the heart from the top to the bottom. The heart contracts and pumps blood due to this electrical signal. The normal rate at which the heart beats in a healthy adult heart at rest is 60 to 100 beats per minute (bpm). It may vary depending on the age, health condition, and activity of an individual. Children usually show faster heartbeat that is considered as normal. As we grow older, it gradually decreases [82].

Pulse Pressure

Pulse pressure (PP) is defined as the difference between systolic and diastolic blood pressure (SBP and DBP). It is the pulsatile component of the total blood pressure wave. There is a steady component as well that is called the mean blood pressure

(MBP). It is the average of systolic and diastolic blood pressure. PP depends on stiffness of the arteries, ventricular ejection, and the timing of wave reflections. Cardiac output and peripheral vascular resistance are the major factors of MBP. In recent studies, it shows PP plays more important role than MBP determining cardiovascular risk in hypertensive patients [76, 83]. An increased pulse pressure is an independent marker for cardiovascular (CV) complications, specially myocardial infarction, congestive heart failure, and other CV deaths [84-86]. It has been known for a while that CV mortality is strongly related to higher systolic pressure [83, 87]. However at any given systolic pressure, CV fatality is higher in case of lower diastolic pressure i.e. higher PP [83]. Two major physiological mechanisms determine pulse pressure. First, higher SBP affects the level of end-systolic stress and enhances the chance of cardiac hypertrophy [78, 85]. Second, lower DBP modifies coronary perfusion and therefore, favors myocardial ischemia [75, 77, 85, 88]. Both cardiac and arterial factors, including ventricular ejection, arterial stiffness, and modifications of the amplitude and timing and site of wave reflections are the determinants of a higher PP [88, 89]. Ventricular ejection remains normal or even lower in older patients. In such cases, alteration of wave reflection and increase of arterial stiffness are the main determinants of higher PP. A 10 mm Hg increase in PP indicates higher risk of major cardiovascular complications keeping the mean pressure and other covariates controlled. The risk associated with the increased PP ranges from 13% for all coronary end points ($P=0.02$) to nearly 20% for cardiovascular mortality ($P=0.01$) [83].

Methods

Data Collection

A fingertip was placed in front of a smart phone camera (Samsung Galaxy S3). Video was recorded for approximately 20 s at a sampling rate of 30 frame/s under ambient light. A standard tool (Omron BP785 IntelliSense Automatic Blood Pressure Monitor) was used to measure and validate the blood pressure and pulse rate readings. All the analysis was done assuming the readings from the standard tool were reasonably accurate. The hand of the subject was kept at rest on a table top during data recording. Subjects were asked to sit and relax for five minutes before every set of data collection. Data was taken alternatively with Omron BP785 and smartphone in order to keep the conditions as similar as possible. Simultaneous data collection, which would be the ideal scenario to calibrate the system accurately, was not possible due to working principle of any sphygmomanometer. It alters the pressure wave in the artery in the process of blood pressure measurement. In order to avoid this problem and keeping the BP unaltered, data was collected separately with Omron BP785 and smartphone but with a minimal gap in time.

Feature Extraction

Once the video was recorded, it was transferred to a standard computer. All the frames from the video were extracted using MATLAB. Each frame was converted into a grayscale image. Intensity of each pixel of a frame was calculated and averaged to find the mean intensity of the frame. Since data was recorded with a specified sampling rate, each frame was associated with a point in time. Once mean intensities for all the frames were computed, time variation of intensity was plotted. Heart rate and pulse pressure were extracted from the intensity plot.

Results and Discussion

Optical Absorption

When light passes through a medium and incidents on an interface (surface between two different media), it changes direction in the second medium. Part of the light bounces back to the first medium, called reflected light and the part passed through the interface into second medium is called refracted light. Some light will be absorbed while travelling through a medium.

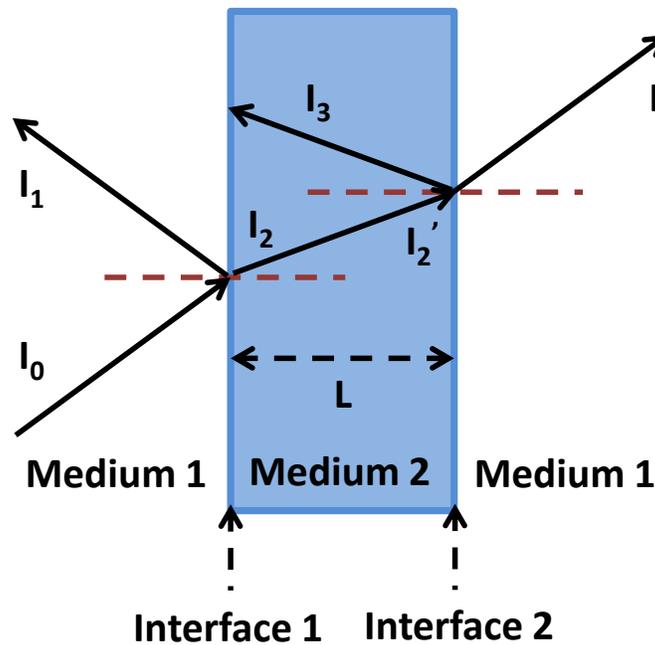


Figure 5-3 A simple model of light passing through a lossy medium (medium 2). Medium 1 is air

$$R_1 = \left(\frac{n_f - n_a}{n_f + n_a} \right)^2 \text{ and } R_2 = \left(\frac{n_a - n_f}{n_a + n_f} \right)^2 \quad (\text{Equation 5-1})$$

n_a and n_f are the refractive indices of air and lossy medium (finger), respectively.

R_1 and R_2 are the reflectances of interface 1 and 2, respectively. From Figure 5-3

$$I_1 = R_1 I_0 \text{ and } I_2 = (1 - R_1) I_0$$

$$I_2' = (1 - R_1)I_0 \exp \left[- \int_0^L \alpha(x) dx \right] \text{ and } I_3 = R_2(1 - R_1)I_0 \exp \left[- \int_0^L \alpha(x) dx \right]$$

$$I = (1 - R_1)(1 - R_2)I_0 \exp \left[- \int_0^L \alpha(x) dx \right] \quad (\text{Equation 5-2})$$

I_0 and I are the intensities of light before and after passing the lossy medium (Medium 2 of Figure 5-3) of L thickness respectively. $\alpha(x)$ is the absorption coefficient at a distance x from interface 1. If n_f and $\alpha(x)$ are assumed as constant over the thickness (L) and $R_1 = R_2$ then Equation 5-2 becomes

$$I = (1 - R_1)^2 I_0 \exp[-\alpha L] \quad (\text{Equation 5-3})$$

If we assume the R_1 is small and thickness of the fingertip is fairly constant, then Equation 5-3 becomes

$$I_d \approx I_0 \exp[-\alpha_d L] \text{ for diastolic phase}$$

$$I_s \approx I_0 \exp[-\alpha_s L] \text{ for systolic phase}$$

Where I_d and I_s are intensity received at the sensor and α_d and α_s are the effective absorption coefficients during diastolic and systolic phase respectively.

During systolic phase, the heart pumps blood out to the organs. As a result, the arteries are engorged with blood in this phase. On the other hand, blood is sucked from organs into the heart during a diastolic phase. Figure 5-4 demonstrates how blood flow changes in a fingertip during systolic and diastolic phases and how that produces intensity variations in the sensor (camera).

The effective absorption coefficient α_s is higher than α_d as systolic phase pushes the blood to the organs. Consequently I_d is higher than I_s . That means less amount of light is absorbed due to the minimal blood presence in the tip during diastolic phase. As a result, it allows more light to go into the camera and a peak in the intensity plot is observed. On the contrary, systolic phase pumps more blood into fingertip and more light is absorbed. A dip in the light intensity at this point was observed. The recording of the

intensity of the transmitted light through the finger showed intensity variation as shown in Figure 5-5 similar to pressure variation of a cardiac cycle.

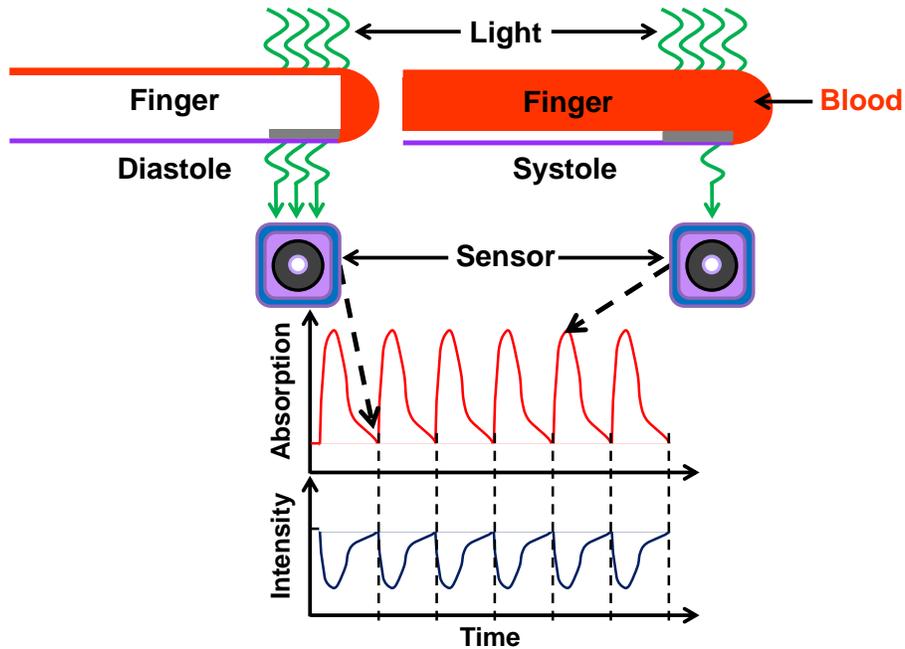


Figure 5-4 Light absorption in fingertip during cardiac cycle

Every downward peak on the intensity plot corresponded to the systolic phase as shown in Figure 5-5. On the rising edge of the intensity dip, there was a point where the intensity fell momentarily and then continues to rise again. This point is called the Dicrotic notch. It is associated with the sudden increase of the pressure in the aorta due to the closure of aortic valve. Diastolic phase starts from Dicrotic notch to the starting of next intensity fall. Figure 5-5 explains systolic, diastolic phase of cardiac cycle and Dicrotic notch.

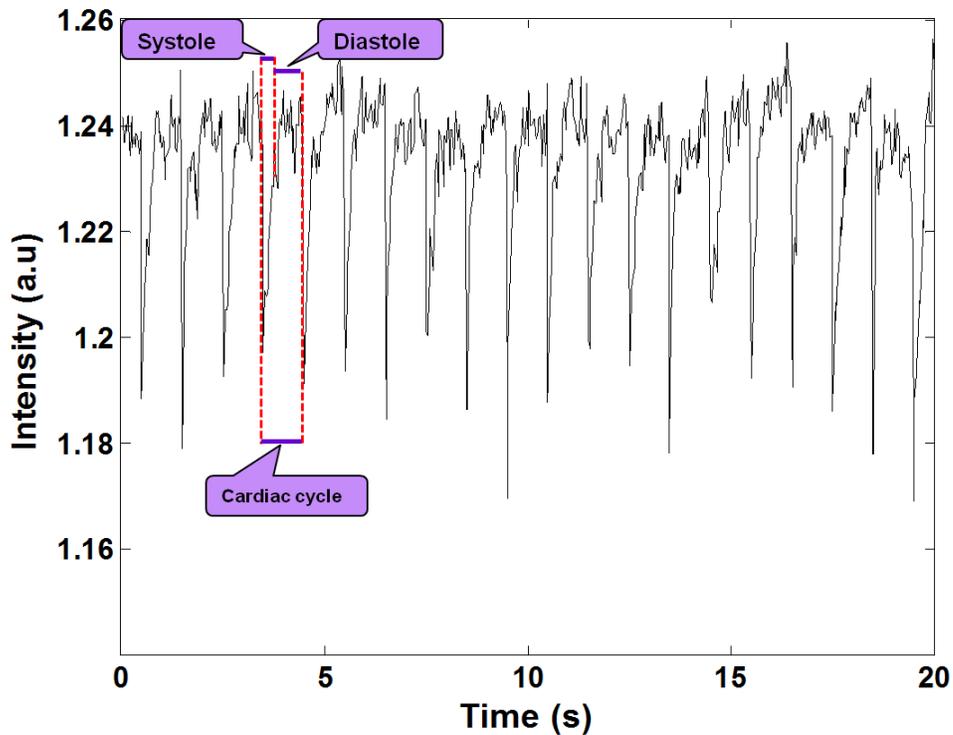


Figure 5-5 Time variation of light intensity on a fingertip during cardiac cycle

Pulse Rate and Pulse Pressure Measurements

In order to compare the results, pulse rate of a subject was measured with a standard BP measuring tool and with a smartphone. Pulse rate was monitored for the same subject and under the same environmental conditions. The results are shown in Figure 5-6. The average pulse rate obtained from automatic BP monitor and smartphone are 67.67 ± 1.92 and 66 ± 3.22 BPM. Since the two results are very close, it indicates that the smartphone method has the potential to be an alternative way of measuring pulse rate in a more convenient way.

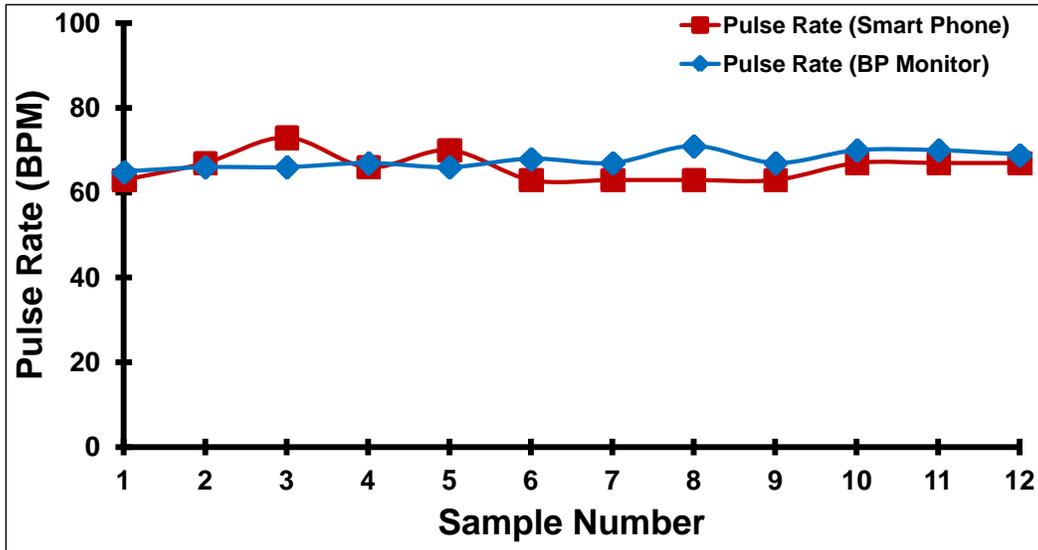


Figure 5-6 Pulse rate measurement using a standard BP monitor and smart phone

Systolic and diastolic pressures were measured using an automatic BP monitor and pulse pressure was then calculated. With a variable light source like ambience, the smartphone method could only measure pulse pressure, not the absolute systolic and diastolic pressure. A comparative picture of pulse pressures is shown in Figure 5-7. The average pulse pressure measured by automatic BP monitor and calibrated smartphone data are 26.67 ± 2.71 and 26.67 ± 2.29 mm Hg. Since pulse pressure can be used as an independent biomarker for cardiovascular diseases and it can be measured with a smartphone with reasonable accuracy, the new system offered a much more convenient way to prevent cardiovascular complications through frequent monitoring and thereby reducing the fatalities due to CV diseases.

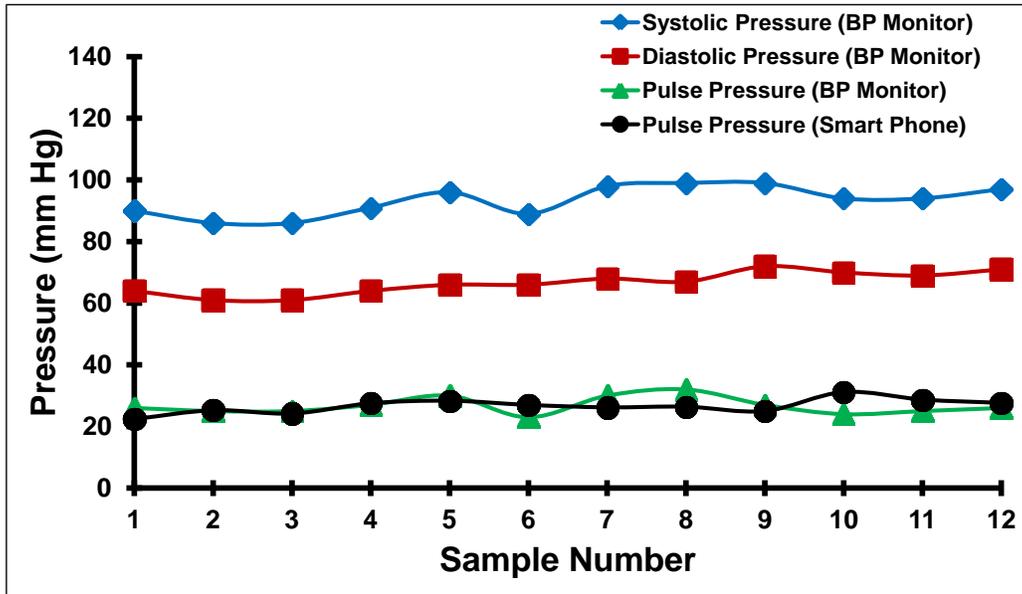


Figure 5-7 Comparison of blood pressure measurement using a standard monitoring device and smartphone

The major advantage of the developed approach is that it offers an easy and convenient way to monitor heart conditions. On top of that, it allows to store this information which is ready to be analyzed by professionals if needs be. Smartphones can store information to a computer/cloud with simple interfaces and allow the patients to share this information to multiple caregivers. If a physician can access the pulse rate and BP information continuously, he/she can help to prevent many CV complications like myocardial infarction, cardiac arrhythmia, hypertension etc. without seeing the patient physically every now and then. Continuous monitoring of BP will also improve the prognosis of other diseases like cerebrovascular diseases, diabetes etc.

Challenges and Solutions

There are a number of sources that can induce noise and fluctuation while collecting data. Since the data was collected in the form of light, a little movement could

induce fluctuation noise. There are always involuntary muscular movements (e.g. from breathing) and these can't be eliminated completely. Multiple set of data needs to be collected from each subject to nullify these effects. Taking average of multiple pulses from each set of data can improve the scenario.

Humans have a wide range of physiological structures depending on the ancestry and demographic location they are brought up. Skin complexions, skin thickness, tissue thickness, the amount of blood, and the amount of hemoglobin carried by each red blood cell are different for each individual. As a result, a single calibration doesn't work well for everybody. In order to solve this problem, a calibration of the system has to be incorporated for different groups of population based on gender, age, skin complexion, and ethnicity. Once the calibration is done, a database of calibration factors will be formed. Each subject will be required to enter these parameters to calibrate the system. Once entered, the system will pull up the appropriate calibration factor for that particular subjects group.

Chapter 6

Summary

- Even with the ever increasing research and findings, cancer is still one of the major causes of death. A diagnostic tool that can hint the patient a little earlier gives the patient a better chance to survive. An aptamer functionalized micropore system was developed to provide a mean to differentiate between cancer and healthy cells. The number of cancer cells present in a sample can also be counted with this method. This estimated number of cancer cells may be a good indicator to determine the next level of therapy. The set-up for single cell differentiation helps to provide more accurate results. This new method has the potential to improve the diagnosis even further. An easily accessible and affordable diagnostic tool can prevent cancers from escalating into more lethal stages. It's vital to have such system within the reach of mass populations.
- Living in an age of internet requires us to keep track of multiple online accounts including our emails, banks, and social media. Each one of them needs a password to authenticate the identity. The number of accounts need to keep track has been significantly grown in the last few years with the increasing connectivity. It requires significant amount of efforts and attentions to put strong passwords to each of these accounts and store them in a safe place. Speaker recognition system developed here may provide a solution to the problem. A unique set of words with the unique voice of each speaker may very well replace the traditional authentication system that uses text passwords. It would be very difficult for a fraud to access someone's account if it's protected with the owner's voice. The same technology can be used to prevent fraudulent access to home and automobiles.

- Even though speech recognition is a fairly old technology, its therapeutic application for children is still at its infancy. There are few applications available on Android and iOS platforms that are targeted to help children to learn and speak new languages. Target for such applications are almost entirely recreational, not therapeutic. None of them has the functionality to estimate how accurately the words are pronounced. The analysis of sounds also provides a path to investigate which part of the sound producing apparatus is not functioning properly as the location of each fundamental sound is already known. Moreover, the current applications don't have feature(s) tracking individual word and overall learning progress. The application developed here provides a method to track and record all words produced by a child undergoing speech therapy and analyzed to determine the appropriate therapy. It can serve as a monitoring tool to speech therapist by supplying objective data. It has the potential to investigate the source location of the speech disorder noninvasively. The progress of each child can be easily quantified and tracked.
- According to Joint Commission of Accreditation of Healthcare Organizations, 65% of hospital death and 55% of medication errors are directly related to faulty communication in US [90]. One of the major reasons for faulty communication between patients and health professionals is accent which makes it difficult for the other party to understand what is spoken. Such miscommunication can be fatal in hospitals and healthcare facilities. For example, if a nurse makes a mistake regarding the dose of a drug due to faulty communication, it would be very dangerous. Having deep accents can be a problem even for the people who are fluent in English. Visual accent trainer provides a preventive solution to this problem. Any individual can use it to get rid of his/her accent. The application developed provides real-time visual feedback on the sound production. It allows a user to try multiple times of a particular

word until the target is met. The application can help people who want learn a foreign language without accent. Due to the globalization, learning a new language without accent can provide an extra edge to any individual. Hearing impaired or deaf people can also be benefitted using this as they can't hear their own sound production. It's difficult for them to adjust any correction in the sound production. By providing visual feedback, VAT can close the audio-visual loop for hearing impaired people. The application can run on all standard computing platforms while leveraging cloud storage.

- The most prevalent cause of deaths in modern society is cardiovascular diseases. Many people still die due to heart related complications each year because there is no convenient way of monitoring blood pressure and heart rate on a regular interval and tracking them for long time. Many of these deaths can be prevented with a reliable and readily accessible heart monitoring tool. Blood pressure and heart rate are the two most vital indicators of many cardiovascular diseases. A noninvasive method to measure these two vital indicators may improve these scenarios. The developed technique has the potential to provide a smartphone application that can be used to monitor and track pulse pressure and heart rate. Doctors can access the patient data through cloud storage without the requirement of the patient physically present in their offices. With the more comprehensive and frequent measurements of pulse pressure and heart rate, doctors may be able to anticipate any heart condition long before its lethal stage.

Future Scope

- In this work, only one type of cancer cell (hGBM) is studied. But the same method can be applied to other type of cancers (lung cancer, bladder cancer, pancreatic cancer, etc.). A micropore/nanopore with the comparable size of the target entity needs to be fabricated along with the synthesis of an aptamer that has specific affinity for the biomarker present on the target cell.
- As mentioned earlier, the condition of sound producing apparatus is also embedded in the speech. One future avenue can be the extraction of features from speech that represents the health conditions of vocal cords or laryngeal tube. It would be a noninvasive way to monitor the health of sound producing apparatus.
- The speech recognition application for children works with only English sound for now, it has the potential to include other languages as well. Incorporating multiple language can be an important future avenue to pursue. Another major future direction might be forming a database with large number of voice samples collected locally. This database would be more useful than the standard database as it might be formed with sounds from a very different group of people.
- The smartphone based heart monitoring system developed here can measure only pulse pressure (difference between SBP and DBP). Even though increased pulse pressure is a good indicator of many cardiovascular diseases including hypertension, it would be very useful if the same method can measure systolic and diastolic pressure separately. A reliable noninvasive way to measure blood pressure using a cell phone would replace the traditional way of measuring blood pressure with bulky tools.

Appendix A

Preparation of Ag/AgCl electrodes

First, the Ag wires were cut into appropriate pieces (5-8 cm). The sized wires were cleaned with ethanol. Then those were immersed into Clorox for 15 minutes. A light gray color emerged on wires that showed the formation of AgCl was done appropriately. Wires were then rinsed with DI water before use. Leaving the electrodes dipped in Clorox solution after the experiment grew the AgCl coating again. As a result, same pair of electrodes can be used multiple times.

Appendix B

Isolation and Culturing of hGBM Cells

Human GBM (hGBM) samples were obtained from University of Texas Southwestern Medical Center (Dallas, TX). Specimen with average volume of more than 50 mm³ was placed into ice-cold HBSS medium immediately after removal from the brain. Red blood cells were removed by Lymphocyte-M (Cedarlane Labs). Then the tumor tissue was gently dissociated with papain (2%) and dispase (2%), triturated, and labeled with a CD133/2 (293C₃)-PE antibody (Miltenyi Biotec) and sorted with FACSCalibur machine (BD Biosciences). A serum-free DMEM/F-12 medium was used as suspension medium for cells consisting of 20 ng/mL of mouse EGF (PeproTech), 20 ng/mL of basic fibroblast growth factor (PeproTech), 1X B27 supplement (Invitrogen), 1X Insulin-Transferrin-Selenium-X (Invitrogen), and 100 units/mL penicillin–100 µg/mL streptomycin (HyClone). Then live cells were plated at a density of 3 X 10⁶ cells/60-mm plate. Both CD133+ and CD133- fractions underwent clonal expansion and formed orthotopic tumors. CD133+ fraction was used as hGBM cells throughout our study. Then hGBM cells were transduced with a lentivirus expressing m-cherry fluorescent protein.

Appendix C
Fast Fourier Transform

FFT stands for fast Fourier transform. It does not refer to a new or different type of Fourier transform rather refers to a very efficient algorithm for computing the discrete Fourier transform (DFT). The time taken to compute a DFT on a computer depends essentially on the number of multiplications involved. N^2 multiplications are required to compute DFT of a data series with N elements. But FFT requires only $N \log_2(N)$. The mathematical insight which leads to this algorithm is the realization that a DFT of a sequence of N points can be written in terms of two DFTs of length $N/2$. If the length of the sequence (N) is a power of two, it is possible to apply this mathematical decomposition recursively until there is only a single point to compute DFT. Even if N is not a power of 2, it can be made so by appending appropriate number of zeros at the end. DFT of a sequence $x(n)$ with length N ,

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-\frac{2\pi jnk}{N}} = \sum_{n=0}^{N-1} x[n] W_N^{nk}$$

where $W_N^{nk} = e^{-\frac{2\pi jnk}{N}}$. It is easy to realize that the same values of W_N^{nk} are calculated many times during the DFT computation. Using the symmetry property the above expression can be split into two similar terms.

$$\begin{aligned} X[k] &= \sum_{n=0}^{N-1} x[n] e^{-\frac{2\pi jnk}{N}} = \sum_{n=0, n \text{ is even}}^{N-1} x[n] W_N^{nk} + \sum_{n=0, n \text{ is odd}}^{N-1} x[n] W_N^{nk} \\ &= \sum_{r=0}^{\frac{N}{2}-1} x[2r] W_N^{2rk} + \sum_{r=0}^{N-1} x[2r+1] W_N^{(2r+1)k} = \sum_{r=0}^{\frac{N}{2}-1} x_1[r] W_{\frac{N}{2}}^{rk} + W_N^k \sum_{r=0}^{N-1} x_2[r] W_{\frac{N}{2}}^{rk} \end{aligned}$$

where $x_1[r] = x[2r]$ and $x_2[r] = x[2r+1]$. So $X[k] = X_1(k) + W_N^k X_2(k)$. From the analysis, it is evident that an N point DFT can be evaluated by computing two $N/2$ points DFT and adding them. This process can be continued until there is one data point left. Due to the exponential nature of the algorithm, it enhances the speed of the computation.

Appendix D
Correlation Coefficient

Linear correlation coefficient is a measure of the strength and the direction of a linear relationship between two variables. The following formula can be used to compute it.

$$r = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2} \sqrt{n(\sum y^2) - (\sum y)^2}}$$

where x and y are two data sets of length n and r is the correlation coefficient between them. It can be used to compute matching between two data sets.

References

- [1] M. Heron, "Deaths: Final Data for 2010," *National Vital Statistics Reports*, vol. 62, pp. 1-97, 2013.
- [2] S. D. Markowitz and M. M. Bertagnoli, "Molecular basis of colorectal cancer," *New England Journal of Medicine*, vol. 361, pp. 2449-2460, 2009.
- [3] A. Kitamura, W. Hosoda, E. Sasaki, T. Mitsudomi, and Y. Yatabe, "Immunohistochemical detection of EGFR mutation using mutation-specific antibodies in lung cancer," *Clinical Cancer Research*, vol. 16, pp. 3349-3355, 2010.
- [4] D. Hanahan and R. A. Weinberg, "Hallmarks of cancer: the next generation," *Cell*, vol. 144, pp. 646-674, 2011.
- [5] P. Anand, A. B. Kunnumakara, C. Sundaram, K. B. Harikumar, S. T. Tharakan, O. S. Lai, B. Sung, and B. B. Aggarwal, "Cancer is a preventable disease that requires major lifestyle changes," *Pharmaceutical Research*, vol. 25, pp. 2097-2116, 2008.
- [6] C. f. D. C. a. Prevention, "Deaths: Final Data for 2013," *National Vital Statistics Reports*, vol. 64, pp. 1-71, 2013.
- [7] A. C. Society, "Cancer Facts & Figures 2014," *Atlanta: American Cancer Society*, 2014.
- [8] H. Mohamed, L. D. McCurdy, D. H. Szarowski, S. Duva, J. N. Turner, and M. Caggana, "Development of a rare cell fractionation device: application for cancer detection," *NanoBioscience, IEEE Transactions on*, vol. 3, pp. 251-256, 2004.
- [9] L. R. Huang, E. C. Cox, R. H. Austin, and J. C. Sturm, "Continuous particle separation through deterministic lateral displacement," *Science*, vol. 304, pp. 987-990, 2004.

- [10] P. R. C. Gascoyne, J. Noshari, T. J. Anderson, and F. F. Becker, "Isolation of rare cells from cell mixtures by dielectrophoresis," *Electrophoresis*, vol. 30, pp. 1388-1398, 2009.
- [11] M. P. MacDonald, S. Neale, L. Paterson, A. Richies, K. Dholakia, and G. C. Spalding, "Cell cytometry with a light touch: sorting microscopic matter with an optical lattice," *Journal of Biological Regulators and Homeostatic Agents*, vol. 18, pp. 200-205, 2004.
- [12] S. Nagrath, L. V. Sequist, S. Maheswaran, D. W. Bell, D. Irimia, L. Ulkus, M. R. Smith, E. L. Kwak, S. Digumarthy, A. Muzikansky, P. Ryan, U. J. Balis, R. G. Tompkins, D. A. Haber, and M. Toner, "Isolation of rare circulating tumour cells in cancer patients by microchip technology," *Nature*, vol. 450, pp. 1235-1239, 2007.
- [13] A. G. J. Tibbe, B. G. de Grooth, J. Greve, G. J. Dolan, C. Rao, and L. W. M. M. Terstappen, "Magnetic field design for selecting and aligning immunomagnetic labeled cells," *Cytometry*, vol. 47, pp. 163-172, 2002.
- [14] W. He, H. Wang, L. C. Hartmann, J.-X. Cheng, and P. S. Low, "In vivo quantitation of rare circulating tumor cells by multiphoton intravital flow cytometry," *Proceedings of the National Academy of Sciences*, vol. 104, pp. 11760-11765, 2007.
- [15] M. Toner and D. Irimia, "Blood-on-a-chip," *Annu. Rev. Biomed. Eng.*, vol. 7, pp. 77-103, 2005.
- [16] I. O. Baas, F. M. V. D. Berg, J.-W. R. Mulder, M. J. Clement, R. J. C. Slebos, S. R. Hamilton, and G. J. A. Offerhaus, "Potential false-positive results with antigen enhancement for immunohistochemistry of the p53 gene product in colorectal neoplasms," *The Journal of Pathology*, vol. 178, pp. 264-267, 1996.

- [17] F. Dalle, J. Lopez, D. Caillot, B. Cuisenier, A. E. Laubriet, L. Dumont, and A. Bonnin, "False-positive results caused by cotton swabs in commercial *Aspergillus* antigen latex agglutination test," *European Journal of Clinical Microbiology & Infectious Diseases*, vol. 21, pp. 130-132, 2002.
- [18] Y. Wan, Y.-t. Kim, N. Li, S. K. Cho, R. Bachoo, A. D. Ellington, and S. M. Iqbal, "Surface-immobilized aptamers for cancer cell isolation and microscopic cytology," *Cancer Research*, vol. 70, pp. 9371-9380, 2010.
- [19] D. H. J. Bunka and P. G. Stockley, "Aptamers come of age - at last," *Nature Reviews Microbiology*, vol. 4, pp. 588-596, 2006.
- [20] B. A. Sullenger and E. Gilboa, "Emerging clinical applications of RNA," *Nature*, vol. 418, pp. 252-258, 2002.
- [21] K. Nagel-Wolfrum, C. Buerger, I. Wittig, K. Butz, F. Hoppe-Seyler, and B. Groner, "The Interaction of Specific Peptide Aptamers With the DNA Binding Domain and the Dimerization Domain of the Transcription Factor Stat3 Inhibits Transactivation and Induces Apoptosis in Tumor Cells," *Molecular Cancer Research*, vol. 2, pp. 170-182, 2004.
- [22] J. Charlton, J. Sennello, and D. Smith, "In vivo imaging of inflammation using an aptamer inhibitor of human neutrophil elastase," *Chemistry & Biology*, vol. 4, pp. 809-816, 1997.
- [23] H. W. Chen, C. D. Medley, K. Sefah, D. Shangguan, Z. Tang, L. Meng, J. E. Smith, and W. Tan, "Molecular Recognition of Small-Cell Lung Cancer Cells Using Aptamers," *ChemMedChem*, vol. 3, pp. 991-1001, 2008.
- [24] O. C. Farokhzad, S. Jon, A. Khademhosseini, T.-N. T. Tran, D. A. LaVan, and R. Langer, "Nanoparticle-aptamer bioconjugates a new approach for targeting prostate cancer cells," *Cancer Research*, vol. 64, pp. 7668-7672, 2004.

- [25] J. A. Phillips, Y. Xu, Z. Xia, Z. H. Fan, and W. Tan, "Enrichment of cancer cells using aptamers immobilized on a microfluidic channel," *Analytical Chemistry*, vol. 81, pp. 1033-1039, 2008.
- [26] W. Asghar, Y. Wan, A. Ilyas, R. Bachoo, Y.-t. Kim, and S. M. Iqbal, "Electrical fingerprinting, 3D profiling and detection of tumor cells with solid-state micropores," *Lab on a Chip*, vol. 12, pp. 2345-2352, 2012.
- [27] Y. Wan, Y. Liu, P. B. Allen, W. Asghar, M. A. I. Mahmood, J. Tan, H. Duhon, Y.-t. Kim, A. Ellington, and S. M. Iqbal, "Capture, Isolation and Release of Cancer Cells with Aptamer-functionalized Glass Beads," *Lab on a Chip*, pp. 4693-4701, 2012.
- [28] Y. Wan, M. Mahmood, N. Li, P. B. Allen, Y.-t. Kim, R. Bachoo, A. D. Ellington, and S. M. Iqbal, "Nanotextured substrates with immobilized aptamers for cancer cell isolation and cytology," *Cancer*, vol. 118, pp. 1145-1154, 2012.
- [29] World Health Organization and Unaid, *Prevention of cardiovascular disease: World Health Organization*, 2007.
- [30] G. S. Roberts, D. Kozak, W. Anderson, M. F. Broom, R. Vogel, and M. Trau, "Tunable nano/micropores for particle detection and discrimination: Scanning ion occlusion spectroscopy," *Small*, vol. 6, pp. 2653-2658, 2010.
- [31] S. Dhar, F. X. Gu, R. Langer, O. C. Farokhzad, and S. J. Lippard, "Targeted delivery of cisplatin to prostate cancer cells by aptamer functionalized Pt (IV) prodrug-PLGA-PEG nanoparticles," *Proceedings of the National Academy of Sciences*, vol. 105, pp. 17356-17361, 2008.
- [32] M. M. Bellah, S. Iqbal, and Y.-T. Kim, "Differential behavior of EGFR-overexpressing cancer cells through aptamer-functionalized micropores," *Microfluidics and Nanofluidics*, vol. 17, pp. 983-992, 2014.

- [33] J. G. M. Klijn, M. P. Look, H. Portengen, J. Alexieva-Figusch, W. L. J. van Putten, and J. A. Foekens, "The prognostic value of epidermal growth factor receptor (EGF-R) in primary breast cancer: results of a 10 year follow-up study," *Breast Cancer Research and Treatment*, vol. 29, pp. 73-83, 1994.
- [34] M. A. Cobleigh, C. L. Vogel, D. Tripathy, N. J. Robert, S. Scholl, L. Fehrenbacher, J. M. Wolter, V. Paton, S. Shak, and G. Lieberman, "Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease," *Journal of Clinical Oncology*, vol. 17, pp. 2639-2639, 1999.
- [35] F. R. Hirsch, M. Varella-Garcia, P. A. Bunn, W. A. Franklin, R. Dziadziuszko, N. Thatcher, A. Chang, P. Parikh, J. R. Pereira, and T. Ciuleanu, "Molecular predictors of outcome with gefitinib in a phase III placebo-controlled study in advanced non-small-cell lung cancer," *Journal of Clinical Oncology*, vol. 24, pp. 5034-5042, 2006.
- [36] G. J. Riely, J. Marks, and W. Pao, "KRAS mutations in non-small cell lung cancer," *Proceedings of the American Thoracic Society*, vol. 6, pp. 201-205, 2009.
- [37] M. S. Brose, P. Volpe, M. Feldman, M. Kumar, I. Rishi, R. Gerrero, E. Einhorn, M. Herlyn, J. Minna, and A. Nicholson, "BRAF and RAS mutations in human lung cancer and melanoma," *Cancer Research*, vol. 62, pp. 6997-7000, 2002.
- [38] P. Lipponen and M. Eskelinen, "Expression of epidermal growth factor receptor in bladder cancer as related to established prognostic factors, oncoprotein (c-erbB-2, p53) expression and long-term prognosis," *British Journal of Cancer*, vol. 69, p. 1120, 1994.

- [39] A. A. Khalil, N. F. Kabapy, S. F. Deraz, and C. Smith, "Heat shock proteins in oncology: diagnostic biomarkers or therapeutic targets?," *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, vol. 1816, pp. 89-104, 2011.
- [40] W. C. Cho, "Contribution of oncoproteomics to cancer biomarker discovery," *Molecular Cancer*, vol. 6, pp. 25-27, 2007.
- [41] W. Asghar, A. Ilyas, R. R. Deshmukh, S. Sumitsawan, R. B. Timmons, and S. M. Iqbal, "Pulsed plasma polymerization for controlling shrinkage and surface composition of nanopores," *Nanotechnology*, vol. 22, pp. 285304-285311, 2011.
- [42] S. E. Osborne, I. Matsumura, and A. D. Ellington, "Aptamers as therapeutic and diagnostic reagents: problems and prospects," *Current Opinion in Chemical Biology*, vol. 1, pp. 5-9, 1997.
- [43] S. M. Iqbal, D. Akin, and R. Bashir, "Solid-state nanopore channels with DNA selectivity," *Nature Nanotechnology*, vol. 2, pp. 243-248, 2007.
- [44] G. Carpenter and S. Cohen, "Epidermal growth factor," *Annual Review of Biochemistry*, vol. 48, pp. 193-216, 1979.
- [45] C. J. Wikstrand, R. E. McLendon, A. H. Friedman, and D. D. Bigner, "Cell surface localization and density of the tumor-associated variant of the epidermal growth factor receptor, EGFRvIII," *Cancer Research*, vol. 57, pp. 4130-4140, 1997.
- [46] T. Uemura, K. Shiozaki, K. Yamaguchi, S. Miyazaki, S. Satomi, K. Kato, H. Sakuraba, and T. Miyagi, "Contribution of sialidase NEU1 to suppression of metastasis of human colon cancer cells through desialylation of integrin β 4," *Oncogene*, vol. 28, pp. 1218-1229, 2009.
- [47] W. C. Chang, D. Liepmann, and L. P. Lee, "A biomimetic method for extracting leukocytes from blood in microfluidic devices," *Microtechnologies in Medicine &*

- Biology 2nd Annual International IEEE-EMB Special Topic Conference on*, pp. 184-188, 2002.
- [48] P. Rose, *Forensic speaker identification*: CRC Press, 2004.
- [49] J. L. Elman and D. Zipser, "Learning the hidden structure of speech," *The Journal of the Acoustical Society of America*, vol. 83, p. 1615, 1988.
- [50] P. D. Templeton and B. J. Guillemin, "Speaker identification based on vowel sounds using neural networks," *The Australasian Speech Science and Technology Association*, pp. 280-285, 1990.
- [51] J.-D. Wu and Y.-J. Tsai, "Speaker identification system using empirical mode decomposition and an artificial neural network," *Expert Systems with Applications*, vol. 38, pp. 6112-6117, 2011.
- [52] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," *Speech and Audio Processing, IEEE Transactions on*, vol. 3, pp. 72-83, 1995.
- [53] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models," *Digital Signal Processing*, vol. 10, pp. 19-41, 2000.
- [54] R. E. Berg and D. G. Stork, *The Physics of Sound*: Prentice-Hall, 1995.
- [55] <http://training.seer.cancer.gov/images/anatomy/respiratory/larynx.jpg>.
- [56] <https://visualsonline.cancer.gov/retrieve.cfm?imageid=4370&dpi=300&fileformat=jpg>.
- [57] R. R. Fay and A. N. Popper, "Evolution of hearing in vertebrates: the inner ears and processing," *Hearing Research*, vol. 149, pp. 1-10, 2000.

- [58] L. U. E. Kohlloffel, "A study of basilar membrane vibrations III. The basilar membrane frequency response curve in the living guinea pig," *Acta Acustica united with Acustica*, vol. 27, pp. 82-89, 1972.
- [59] L. Robles, M. A. Ruggero, and N. C. Rich, "Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases," *The Journal of the Acoustical Society of America*, vol. 80, p. 1364, 1986.
- [60] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*: MIT press, 1994.
- [61] B. Schneider, J. Wendler, and W. Seidner, "The Relevance of Stroboscopy in Functional Dysphonias1," *Folia Phoniatrica etLogopaedica*, vol. 54, pp. 44-54, 2002.
- [62] J. G. Å vec and H. K. Schutte, "Videokymography: high-speed line scanning of vocal fold vibration," *Journal of Voice*, vol. 10, pp. 201-205, 1996.
- [63] S. S. Stevens, J. Volkman, and E. B. Newman, "A scale for the measurement of the psychological magnitude pitch," *The Journal of the Acoustical Society of America*, vol. 8, pp. 185-190, 1937.
- [64] A. H. Benade, "Fundamentals of musical acoustics," *Oxford University Press, London.*, 1976.
- [65] G. A. Mensah and D. W. Brown, "An overview of cardiovascular disease burden in the United States," *Health Affairs*, vol. 26, pp. 38-48, 2007.
- [66] "<http://nuclearcardiologyseminars.com/electrocardiography/normal/>."
- [67] M. C. Sanguinetti and M. Tristani-Firouzi, "hERG potassium channels and cardiac arrhythmia," *Nature*, vol. 440, pp. 463-469, 2006.

- [68] A. V. Chobanian, G. L. Bakris, H. R. Black, W. C.ushman, L. A. Green, J. L. Izzo Jr, D. W. Jones, B. J. Materson, S. Oparil, and J. T. Wright Jr, "The seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure: the JNC 7 report," *Jama*, vol. 289, pp. 2560-2571, 2003.
- [69] T. Pickering, "Recommendations for the use of home (self) and ambulatory blood pressure monitoring," *American Journal of Hypertension*, vol. 9, pp. 1-11, 1996.
- [70] J. C. Dekkers, H. Snieder, E. J. C. G. van den Oord, and F. A. Treiber, "Moderators of blood pressure development from childhood to adulthood: a 10-year longitudinal study," *The Journal of Pediatrics*, vol. 141, pp. 770-779, 2002.
- [71] D. A. Lane and P. Gill, "Ethnicity and tracking blood pressure in children," *Journal of Human Hypertension*, vol. 18, pp. 223-228, 2004.
- [72] W. Bao, S. A. Threefoot, S. R. Srinivasan, and G. S. Berenson, "Essential hypertension predicted by tracking of elevated blood pressure from childhood to adulthood: the Bogalusa Heart Study," *American Journal of Hypertension*, vol. 8, pp. 657-665, 1995.
- [73] A. Chiolerio, F. Cachat, M. Burnier, F. Paccaud, and P. Bovet, "Prevalence of hypertension in schoolchildren based on repeated measurements and association with overweight," *Journal of Hypertension*, vol. 25, pp. 2209-2217, 2007.
- [74] R. Din-Dzietham, Y. Liu, M.-V. Bielo, and F. Shamsa, "High blood pressure trends in children and adolescents in national surveys, 1963 to 2002," *Circulation*, vol. 116, pp. 1488-1496, 2007.

- [75] S. S. Franklin, S. A. Khan, N. D. Wong, M. G. Larson, and D. Levy, "Is pulse pressure useful in predicting risk for coronary heart disease? The Framingham Heart Study," *Circulation*, vol. 100, pp. 354-360, 1999.
- [76] R. Asmar, A. Rudnichi, J. Blacher, G. r. M. London, and M. E. Safar, "Pulse pressure and aortic pulse wave are markers of cardiovascular risk in hypertensive populations," *American Journal of Hypertension*, vol. 14, pp. 91-97, 2001.
- [77] G. F. Mitchell, L. A. Moyé, E. Braunwald, J.-L. Rouleau, V. Bernstein, E. M. Geltman, G. C. Flaker, M. A. Pfeffer, and I. Save, "Sphygmomanometrically determined pulse pressure is a powerful independent predictor of recurrent events after myocardial infarction in patients with impaired left ventricular function," *Circulation*, vol. 96, pp. 4254-4260, 1997.
- [78] B. Darne, X. Girerd, M. Safar, F. Cambien, and L. Guize, "Pulsatile versus steady component of blood pressure: a cross-sectional analysis and a prospective analysis on cardiovascular mortality," *Hypertension*, vol. 13, pp. 392-400, 1989.
- [79] "http://cnx.org/contents/22c3f969-ab9d-4df0-8baf-5b1a99372fac@3/Cardiac_Cycle."
- [80] <http://www.nhlbi.nih.gov/health/health-topics/topics/hhw/electrical>.
- [81] <http://www.pearsonhighered.com/marieb-10e-info/assets/pdf/marieb-0321927028-chapter18.pdf>.
- [82] "Your Heart's Electrical System," *National Institutes of Health. National Heart, Lung, and Blood Institute*, 2011.
- [83] J. Blacher, J. A. Staessen, X. Girerd, J. Gasowski, L. Thijs, L. Liu, J. G. Wang, R. H. Fagard, and M. E. Safar, "Pulse pressure not mean pressure determines

- cardiovascular risk in older hypertensive patients," *Archives of Internal Medicine*, vol. 160, pp. 1085-1089, 2000.
- [84] J. A. Millar, A. F. Lever, and V. Burke, "Pulse pressure as a risk factor for cardiovascular events in the MRC Mild Hypertension Trial," *Journal of Hypertension*, vol. 17, pp. 1065-1072, 1999.
- [85] A. Benetos, M. Safar, A. Rudnichi, H. Smulyan, J.-L. Richard, P. Ducimetière, and L. Guize, "Pulse pressure a predictor of long-term cardiovascular mortality in a French male population," *Hypertension*, vol. 30, pp. 1410-1415, 1997.
- [86] A. Benetos, A. Rudnichi, M. Safar, and L. Guize, "Pulse pressure and cardiovascular mortality in normotensive and hypertensive subjects," *Hypertension*, vol. 32, pp. 560-564, 1998.
- [87] A. Sagie, M. G. Larson, and D. Levy, "The natural history of borderline isolated systolic hypertension," *New England Journal of Medicine*, vol. 329, pp. 1912-1917, 1993.
- [88] W. Nichols, M. O'Rourke, and C. Vlachopoulos, *McDonald's blood flow in arteries: theoretical, experimental and clinical principles*: CRC Press, 2011.
- [89] M. E. Safar, B. I. Levy, and H. Struijker-Boudier, "Current perspectives on arterial stiffness and pulse pressure in hypertension and cardiovascular diseases," *Circulation*, vol. 107, pp. 2864-2869, 2003.
- [90] "What Did the Doctor Say?" Improving Health Literacy to Protect Patient Safety," *The Joint Commission of Accreditation of Healthcare Organizations*, p. 48, 2007.

Biographical Information

Md. Motasim Bellah was born and raised in Jessore, a city located on the southwestern part of Bangladesh. He earned Bachelor of Science degree in Electrical and Electronic Engineering from Bangladesh University of Engineering and Technology in January 2008. He then joined LM Ericsson Bangladesh Ltd. Dhaka, Bangladesh as a service engineer in 2008. Later in that year he joined United International University as a lecturer. He taught and developed eight undergraduate courses on circuit analysis, electronics, properties of materials, semiconductor device theory, and electromagnetics. He supervised three undergraduate projects. He moved to USA in January 2010 and started his PhD in the department of electrical engineering at the University of Texas at Arlington. He then joined Nano Bio Lab, a research lab that primarily focuses on different modalities of cancer detection, as a research associate under the supervision of Prof. Samir Iqbal.

His research was primarily focused on developing solid-state micropore based biosensors for early cancer detection. He has also worked on building a smartphone based diagnostic platform for measuring blood pressure and heart rate noninvasively. He has built a speaker recognition system that uses English vowels to identify a person. He has also developed an accent trainer for new language learner that provides real-time visual and quantitative feedback about the sound production, contrary to the traditional accent trainer which plays the sound and asks the learner to mimic it without giving any quantitative feedback. He designed an android application that helps children with speech impediment to learn quickly.