SNAPPY TITLE GOES HERE

A Dissertation
Submitted to
the Temple University Graduate Board

in Partial Fulfillment
of the Requirements for the Degree
DOCTOR OF PHILOSOPHY

by
Christian Radcliffe Ward
May / August / December, 20XX

Examining Committee Members:

Dr. Your Advisor , Advisor, Dept. of Electrial and Computer Engineering
Dr. Member One, Dept. of Z and X
Dr. Member Two, Dept. of Z and X
Dr. Member Three, Dept. of Z and X
Dr. Member Four, External Reader, Dept. of Z and X

ABSTRACT

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque

tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

ACKNOWLEDGEMENTS

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Words.

TABLE OF CONTENTS

APPENDICES

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADHD attention-deficit/hyperactivity disorder

BCI brain-computer interface

CEP Cepstral Coefficient

COH spectral coherence

CRR Correct Recognition Rate

DET Detection Error Tradeoff

EC Eyes Closed

EDF European Data Format

EEG electroencephalogram

EER equal error rate

EM expectation maximiation

EO Eyes Opened

FAR false acceptance rate

FRR false rejection rate

FFT Fast Fourier Transform

GMM-UBM Gaussian Mixture Model-Universal Background Model

GMMHMM Gaussian Mixture Model based Hidden Markov Model

HTER half total error rate

HTK Hidden Markov Toolkit

I-Vector Identity Vector

LDA Linear Discriminate Analysis

MAP maximum a priori

MD Mahalanobis Distance

MFCC Mel Frequency Cepstral Coefficient

MSR Microsoft Research

NEDC Neural Engineering Data Consortium

PhysioNet Database PhysioNet EEG Motor Movement/Imagery Database

PSD Power Spectral Density

RA1 Research Aim 1

RA2 Research Aim 2

RA3 Research Aim 3

SPMD Single Program Multiple Data

TCP Trans-Cranial Parasagittal

TVM total variability matrix

TUH Corpus Temple University EEG Corpus

UBM Universial Background Model

Chapter 1

INTRODUCTION

Chapter 2

BACKGROUND

Chapter 3

METHODS

Those who fail to plan, plan to fail.

Attributed to Benjamin Franklin

The application of Identity Vectors (I-Vectors) on electroencephalograms (EEGs) is a novel concept given that I-Vectors were designed for speech processing. Therefore, there is minimal guidance on how to use I-Vectors on EEG data. As indicated in the background, the foundations of the experiments proposed here came from following the development of I-Vectors within the speech community. The two fields are related in terms of their signal analysis goals, and subject and condition discrimination (Research Aim 1), but their optimization processes may be different (Research Aim 2). In both Aims, the desired goals afford insight into the classification process, which in turn is leveraged into insight about the features, datasets, and EEGs themselves.

3.1 Experimental Outline

The ultimate goal of this research is to provide subject and condition discrimination of EEGs. Prior to this work, this goal was not possible using I-Vectors given the lack of a software tools specifically for EEGs. The first experiments provided classification performance showing that I-Vectors met or exceeded performance of equivalent techniques. Providing competitive classification required an understanding of the technique's trade-offs in terms of features, datasets, and parameters. Running experiments to sweep through the features, datasets, and parameters provided operational thresholds for the datasets, Universal Background Models (UBMs), and UBMs for using I-Vectors based classification on EEGs.

In this work the experiments are classified as *Algorithm Benchmarks*, *Parameter Sweeps*, and *UBM-total variability matrix (TVM) Relationship*. The Algorithm Benchmarks addressed Research Aim 1 (RA1) by testing the performance of I-Vectors against benchmark classifiers, specifically Mahalanobis distance and Gaussian Mixture Model-Universal Background Model (GMM-UBM). The initial comparisons were carried out using parameters borrowed from speech recognition, which then required optimization through the Parameter Sweeps that addressed Research Aim 2 (RA2). Using the optimal classification parameters, the mechansims by which I-Vectors carried out their classification was resolved through analysis of the relationships between the UBMs, TVMs, and feature sets. These UBM-TVM Relationship experiments addressed Research Aim 3 (RA3) and represented the major contribution to understanding EEGs and multi-modal signal analysis.

Each experiment operated on the same fundamental features, datasets, and evaluations as they built upon each other. This chapter details all the components used to build out the experiments. The ensuing three chapters organize present each of the experiments: Chapter 5 - Parameter Sweeps, Chapter 6 - Algorithm Benchmarks, and Chapter 7 - UBM-TVM Relationship.

3.2 Data

Using heterogeneous data is necessary for validating any statistically rigorous method such as I-Vectors, but EEG data is difficult to obtain. Typically, new data is generated as part of research experiments and/or acquired from hospitals, but rarely if ever enters the public domain. This limits innovation to specific combinations of data and techniques. To mitigate this, only the publicly available datasets from PhysioNet EEG Motor Movement/Imagery Database (PhysioNet Database)[1] and Temple University EEG Corpus (TUH Corpus)[2] were used in this work. While not comprehensive in terms of the variety of subjects and conditions used in other studies this collection provided the necessary breadth to validate the goals of this work. These data include EEG from imagined and actual hand, arm, and foot motion, and normal, abnormal, and seizure clinical EEGs from over 600 subjects.

3.2.1 PhysioNet Database

This EEG data comes from the New York State Department of Health's Wadsworth Center [3] and is a component of the PhysioBank archive maintained by MIT's Lab for Computational Physiology[1]. Within the data bank are EEG recordings pertaining to resting states, imagined motion, and motion tasks. The ~~collected~~ data consist of 64 channel EEGs from 109 subjects performing 14 trials: 12 motion and 2 resting calibration (see Figure 3.1). Information about the subjects (age, gender, handedness, etc) is not provided, making subjects and trials the most applicable decision surfaces.

Each 2-minute imagined-motion/motion trial consists of a series of 30 4.1 second tasks. These alternate between rest states and the computer prompted tasks (T1-T4). The tasks consist of opening/closing left or right fist (T1), imagine opening/closing left or right fist (T2), opening/closing both fists or feet (T3), and imagine opening/closing both fists or feet (T4). The two resting state trials, TR1 Eyes Opened (EO) and TR2 Eyes Closed (EC), are one minute recordings of unprompted subject recordings. From this, three dataset

1. **Physio Full** - All fourteen trials (TR01-TR14)

[1] https://www.physionet.org/pn4/eegmmidb/

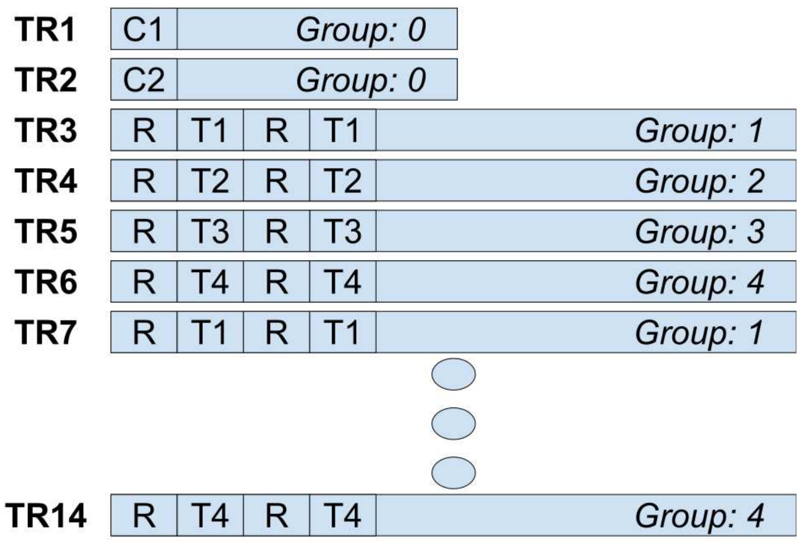

Figure 3.1. Format of PhysioNet Trials. The 12 PhysioNet EEG Motor Movement/Imagery Database (PhysioNet Database) trials are broken down by their tasks into 2 resting trials (TR1, TR2) and the four imagined-motion/motion trials (TR3-TR14). This provided depth for subject evaluation, by adding trials, but also for within subject trial evaluation as the repeated trials could be grouped.

2. **Physio Single** - One trial of each type (TR01-TR06)

3. **Physio Motion** - One of each motion trial (TR03-TR06)

These datasets allowed classification experiments on distinct levels of the data. The highest level was subject classification across trials. Beneath that was subject-trial classification, dependent on matching the correct subject and trial. Finally, within-subject trial classification was possible given the grouping of the repeated trials.

The recordings consist of 64 electrodes sampled at 160Hz following a standard 10-20 layout. A 65th channel provides labels for each task during the trials. Since its introduction in 2009, the PhysioNet Database has been used in biometric classifications [4] with respect to task sensitivity [5], subject independence [6], various subject classification schemes [7, 8], and attempts at content based retrieval [9].

3.2.2 TUH Corpus

The Temple University EEG Corpus (TUH Corpus) contains over 25,000 EEGs with their associated medical evaluations. All data comes from patients seen by Temple University Hospital in Philadelphia, Pennsylvania [2]. These recordings represent considerable breadth and depth in terms of patients, medical conditions, and recording conditions. Seizures were the most common diagnosis for patient's with medical records, but stroke and concussion patients are represented as well, while the majority of all recordings are simply indeterminate. In addition to these patients, there are subsets consisting of normal patients and those with indeterminate conditions considered abnormal. These latter classifications (abnormal/normal) along with seizure patients were used to organize 3 distinct datasets:

1. **TUH Normal** - 50 normal patient sessions
2. **TUH Abnormal** - 50 abnormal patient sessions
3. **TUH Seizure** - 411 seizure patient sessions

These datasets allowed for two types of classification experiments. The first was on the subject level, as each was built from unique subjects. The second was developed by combining the datasets to classify them based upon their condition, abnormal/normal/seizure. Further analysis was possible given the associated medical reports, but beyond the time and scope of this research.

Unlike the PhysioNet Database, the TUH Corpus is *in vivo*, leading to a wide array of recording variation. The electrode configurations, sampling rates, and session counts are at the discretion of medical professionals and not a structured research protocol. As addressed in its public release [2], the most common recording configuration consists of 31 electrodes at a 250Hz sample rate. This is substantially fewer

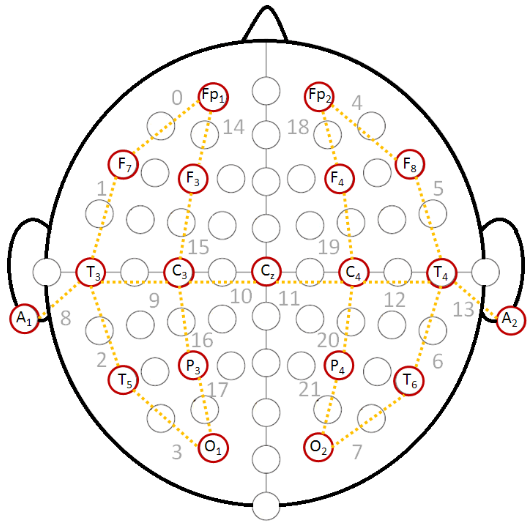

Figure 3.2. Layout of TCP montage for CEP features.. The Trans-Cranial Parasagittal (TCP) montage uses a rostral to caudal differential between electrodes to produce channel data. This differential is applied from the ears inward as well to produce 22 distinct channels. Common electrode names are provide with intermediate electrodes left blank. The gray numbers represent the channel index found in the Temple University EEG Corpus (TUH Corpus).

electrodes than the PhysioNet Database, but is enough to produce clinically common EEG montages[2].

3.2.3 Synthetic Dataset

Developing and testing on experimental data alone would make it impossible to provide validation of the software's efficacy; therefore, a synthetic dataset was built. This controlled dataset allowed for two 'ideal' configurations: (1) a dataset with a common feature across all subjects and (2) a dataset with an unique feature for each subject. These datasets were labeled as *simulated*, *static* (simulated with an additional common feature across subjects) , and *unique* (simulated with a unique feature for each

[2]ACNS - Guideline 3: `http://www.acns.org/UserFiles/file/EEGGuideline3Montage.pdf`

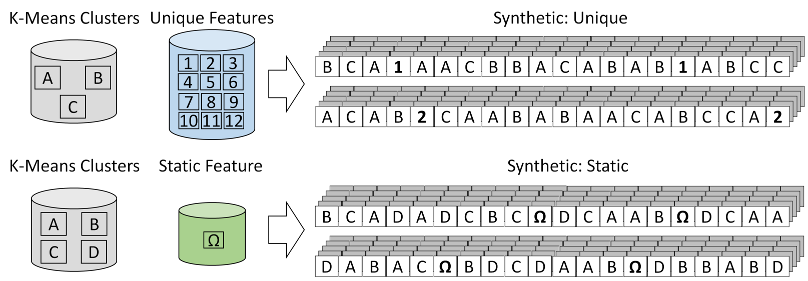

Figure 3.3. Generation of synthetic data from the TUH Corpus.. The GMMHMM modeled data (gray) and the unique (blue) or static (green) features enable the creation of unique and static synthetic data sets. Only 10% of the simulated data is replaced by the external PhysioNet Database feature. The modeling produced features for each epoch's 22 channels simultaneously to keep the channel-epochs temporal synchronized for each of the 12 simulated TUH Corpus subjects.

subject). Each one contained 10 minutes of data for the simulated 12 subjects and their 22 channels, matching the number of channels in the AutoEEG dataset.

Production of the synthetic datasets relied on a Gaussian Mixture Model based Hidden Markov Model (GMMHMM) consisting of 3, 4, or 5 Gaussian models drawn from UBMs. The baseline UBM came from 12 TUH Corpus AutoEEG V1.1.0 subjects using a 16-mixture UBM. The common and unique features came from a single random subject in the PhysioNet Database, also using a 16-mixture UBM. Simulated data contained either 3 or 4 mixtures, allowing the static and unique to add an additional feature containing 4 or 5 mixtures depicted by Figure 3.3.

This produced six unique synthetic data sets: Sim3, Sim4, Sta3, Sta4, Uni3, Uni4, outlined in Figure 3.3. Data was generated for each one-second epoch of each channel as Cepstral Coefficient (CEP) features directly. The distribution of the simulated data followed the weighting of the initial 16 mixture UBM. When the static and unique features were added they overwrote 10% of the simulated data with the new PhysioNet Database-based feature. Authenticity of the raw data was preserved by keeping the synthetic data as similar to the TUH Corpus AutoEEG V1.1.0 dataset as possible, highlighted in Table 3.1.

Table 3.1. Composition of Synthetic Data Sets. Composition of Synthetic Data Sets

Name	Type	Features	Channels	Sampling Rate (Hz)	Duration (s)
AutoEEG	Real	∞	22	100	1200
PhysioNet	Real	∞	64	160	120
Sim3	Simulated	3	22	100	600
Sta3	Static	4	22	100	600
Uni3	Unique	4	22	100	600
Sim4	Simulated	4	22	100	600
Sta4	Static	5	22	100	600
Uni4	Unique	5	22	100	600

3.2.4 Feature Sets

In addition to using multiple datasets, three feature sets were applied to the PhysioNet Database and TUH Corpus: Cepstral Coefficient (CEP), spectral coherence (COH), and Power Spectral Density (PSD). Using multiple feature sets was important because there is no consensus on an optimal feature set for EEGs. PSD features have a long history of use with EEGs [10, 11, 12], as do COH features [13, 14]. CEP are well-established features in the speech processing domain [15, 16]; their application to EEG research was introduced by the Neural Engineering Data Consortium (NEDC) [17].

The COH and PSD features were computed according to the work of LaRocca [13]. The CEP features were built following the standards developed by the speech community [17] and their channels modified to conform with a TCP montage used by neurologists [18]. Thus the feature sets are distinct not only in their mathematical construction, but also their topographical configurations, Table 3.2.

Table 3.2. Feature Set Configurations.

Feature Set Configurations

Name	Type	Features	Channels
CEP	Original	26	22
	Slim	26	22
PSD	Original	40	56
	Slim	40	19
COH	Original	40	1540
	Slim	40	22

As discussed in the background, EEG recordings can use a variety of electrode configurations. For example, the PhysioNet Database contains 64 electrodes of data, while the TUH Corpus contained a myriad of electrode configurations. Therefore the TUH Corpus set was aligned with the most common standard, the TCP montage, resulting in 19 electrodes organized as 22 differential channels. La Rocca's features consisted of 56 PSD channels and 1540 COH channels making for a larger disparity in channels for each feature set. To address this channel imbalance, the TUH Corpus configuration layout was replicated for the PSD and COH feature sets producing two groups of features. The first was the 55 electrode layouts used by La Rocca [13] and the second time was a mirror of the 19 electrodes from the TUH Corpus TCP montage.

This resulted in a slim feature set consistent of the 22 channel CEP, 19 channel PSD, and 22 channel COH. The CEP and COH confirmed to the TCP layout, but the PSD were not converted to keep them as distinct from the COH features as possible.

The benchmark testing against La Rocca's worked used the full feature sets, while all Algorithm Benchmarks and UBM-TVM Relationship experiments used the slim feature sets.

3.2.4.1 Cepstral Features

The CEP-based features were predicated on the success of similar Mel Frequency Cepstral Coefficient (MFCC) used in speech recognition. Their adoption for EEG required shifting from a log frequency scale to linear frequency and adjusting the time windows for the Δ and $\Delta\Delta$ differentials. Generation of these features was introduced and detailed by Harati et al in [17], but is outlined here.

The base feature vector consisted of of nine coefficients (seven cepstral coefficients, the frequency domain energy, and the differential energy). The filter banks actually produce eight spectral coefficients covering the following frequency ranges: $\{0$, 1-10, 11-20, 21-30, 31-40, 41-50, 51-60, 71-80 Hz$\}$. However, the zeroth coefficient is discarded and replaced with the frequency domain energy; the differential frequency energy becomes the ninth term. These filters provided a single energy value after bandpass filtering (Hamming) the Fast Fourier Transform (FFT) for each of the listed frequency ranges.

The two energy terms: frequency domain (E_f) and differential frequency energy (E_d) are given as:

$$E_f = \log\left(\sum_{k=0}^{N-1} |X(k)|^2\right) \tag{3.2-1}$$

$$E_d = \max\Big(E_f(m)\Big) - \min\Big(E_f(m)\Big) \tag{3.2-2}$$

E_f was derived from the outputs of the filter banks where N are the number of filters and X is the filtered cepstrum frequency output. Using these values within the prescribed 0.9s window of samples, the E_d is found by comparing the maximum and minimum E_f values over the range of m elements in the signal window. These built the first nine features with the remaining 17 coming from the first derivative (Δ) and second derivative ($\Delta\Delta$).

The Δ and $\Delta\Delta$ features used the same equations, but with different window sizes:

$$d_t = \frac{\sum_{n=1}^{N} \left[c_{t+n} - -c_{t-n}\right]}{2\sum_{n=1}^{N} n^2} \tag{3.2-3}$$

Here each sample n in the window N was used to produce a derivative for a given coefficient c centered around time t. Zero padding was used to pad the vector near the beginning and ending of the data. The first derivative Δ used $N = 0.9$. Once resolved, the second derivative $\Delta\Delta$ used the Δ values with a new window of $N = 0.3$. In Harati's work [17], the optimal configuration was found to be a 26 feature vector where the $\Delta\Delta$ for E_d was excluded. This configuration was adopted by the research group and became the consistent feature for the experiments in this work.

3.2.4.2 Power Spectral Density Features

PSD features are derived from the sum of energy over a frequency range for a given time sample. Variation in their creation can be found in their frequency range, number of FFT samples, and filtering of the time signal. The variation of PSD based features used in this work are identical to those of La Rocca et al. [13] which used a frequency range of 0-100Hz, a 100-point FFT, and Hanning windows for filters. The final features were 10-second epochs with 40 PSD values evenly spanning 1-40Hz.

The time series data was filtered with 1 second Hanning window using a 0.5 second overlap. This produced 20 filtered samples for each 10 second epoch centered around each 1 second interval from 0 to 9.5 seconds. These filtered samples were evaluated

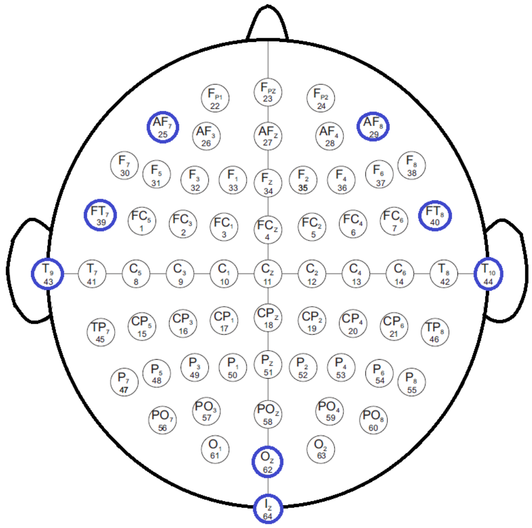

Figure 3.4. Layout of La Rocca's PSD and COH Channels.. The channel layout La Rocca et al used removed 8, highlighted in blue, channels from the overal 64 channel configuration of the PhysioNet Database.

using Welch's averaged modified periodgram (built into Matlab) with a 100 point FFT to produce 1Hz resolution over the range of 0 to 100Hz. La Rocca's work used the PhysioNet Database data which first had to be resampled from 160Hz to 100Hz prior to the filtering.

The resultant 100 energy levels were reduced down to only those spanning 1-40Hz. This reduction in frequencies is necessary given (a) the resampling and (b) that the EEG oscillations of interest Delta (0.5-4Hz), Theta (4-7Hz), Alpha (8-14Hz), Beta (15-29Hz), and Gamma (30-40Hz) fall within that range. This resulted in 40 features per EEG channel. The channel count was reduced to 56 from PhysioNet Database's original 64. The discarded channels, highlighted in Figure 3.4, were AF_7, AF_8, FT_7, FT_8, T_9, T_{10}, O_Z, and I_Z.

While originally designed with the PhysioNet Database in mind, these features were readily adapted to the TUH Corpus. Recordings were resampled to 100Hz and pared down to the match the abbreviated 56 channel layout.

3.2.4.3 Spectral Coherence Features

The COH features were proposed by La Rocca as an improvement over PSD features for subject classification. Measuring coherence between electrodes had been used prior for distinguishing attention-deficit/hyperactivity disorder (ADHD) [19], a general connectivity measure of the brain [20] and auditory oddball paradigms for brain-computer interface (BCI)/P300 responses [21]. Thus they were not novel features, but applied to a broad range of applications beyond subject classification.

These features were generated by quantifying the amount of synchronous energy at each frequency band of each electrode. This was achieved by first building the PSD features and then using them to generate a COH value for each frequency f between two different electrodes i and j, outlined as follows:

$$\mathrm{COH}_{i,j}(f) = \frac{|S_{i,j}(f)|^2}{S_{i,i}(f) \cdot S_{j,j}(f)} \tag{3.2-4}$$

The resultant values were scaled by arctan to normalize their distribution making them bounded on the range $(0, \frac{\pi}{2})$. This configured the final feature set as 1540 'channel' which La Rocca called elements. Each with the 40 distinct frequency bins found through the PSD feature process.

3.2.4.4 Aggregated Datasets

The UBM-TVM Relationship experiments needed subject and condition variation to test classification performance. To achieve, this aggregated datasets were built by combining the PhysioNet Database and TUH Corpus datasets. The combinations of PhysioNet Database's motion data and the TUH Corpus's normal, TUH Corpus's abnormal and normal, or TUH Corpus abnormal, normal, and seizure datasets allowed classification of subjects and known characteristics within a single experiment. This was important to address algorithm robustness and to mitigate any benefits conferred based upon a given dataset-feature-algorithm combination. Each combination was given a designation, Table 3.3 to streamline documentation and discussion.

Table 3.3. Combine Dataset Designations. Combine Dataset Designations

Designation	Dataset 1	Dataset 2	Dataset 3
`AbnNrm`	TUH Abnormal	TUH Normal	-
`AbnSzr`	TUH Abnormal	TUH Seizure	-
`NrmSzr`	TUH Normal	TUH Seizure	-
`AbnMot`	TUH Abnormal	Physio Motion	-
`NrmMot`	TUH Normal	Physio Motion	-
`SzrMot`	TUH Seizure	Physio Motion	-
`AbnNrmSzr`	TUH Abnormal	TUH Normal	TUH Seizure
`AbnNrmMot`	TUH Abnormal	TUH Normal	Physio Motion
`NrmSzrMot`	TUH Normal	TUH Seizure	Physio Motion
`AbnSzrMot`	TUH Abnormal	TUH Seizure	Physio Motion

3.3 Evaluation Metrics

All experiments were run as subject verification tests. This was inline with La Rocca's experiment which used Correct Recognition Rate (CRR) as their sole evaluation metric. However, given the depth of the datasets and parameter testing to be conducted it was necessary to also include the equal error rate (EER) as well. The performance of I-Vectors has typically been reported in terms of EER, while the EEG research community is typically more broadly focused more on CRR. Exceptions in the literature

[5, 22, 23] show results in terms of EER, false acceptance rate (FAR), false rejection rate (FRR), half total error rate (HTER), or Detection Error Tradeoff (DET) curves. For the purposes of this research results were reported in terms of CRR and EER to facilitate readers from both the I-Vector and EEG communities being able to contextualize the experiment performances.

In this work CRR was calculated based on the testing data correctly matching into the enrollment data. The EER was calculated over the entire distance matrix ensuring it evaluated the strength of all matches. This meant if subject 100s second best score was stronger than subject 4's score the EER would be none zero. This is why it was critical to include it for the parameter sweeps, as the CRR masked the majority of the nuance of the full system.

Even with the importance of both metrics, the intended parameter sweeps and comparison points made always displaying both CRR and EER cumbersome and ineffective to the end goal of comparative performance. AS such, the C Metric was defined which combined the CRR and EER by subtracting the EER from the CRR. Thus the threshold for an acceptable C Metric score was set at 0.75 which could represent a CRR of 85% and an EER of 10%. This was primarily used for the expansive Algorithm Benchmarks to showcase performance differences between GMM-UBMs, Mahalanobis Distance (MD), and I-Vectors.

3.3.1 Mixture Size

For UBMs, TVMs, and I-Vectors the dimension of the underlying mixture model is a critical parameter than can affect performance. Effectively, the n dimensional feature space is modeled by m gaussians; these gaussians are used to train the I-Vectors. As has been the case in the speech community [24, 25, 26], it was necessary to determine the size of the mixture model that would optimize I-Vector performance under different circumstances. While some experiments applied GMM-UBMs previously, their protocols and datasets were not a sufficient starting point[27, 23].

These experiments were used to inform the initial mixture sweep range {2, 4, 8, 16. 32, 64, 128, 256, 512, 1024} used as part of the Parameter Sweeps. After which it was expanded to {2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048} for the Algorithm Benchmarks and UBM-TVM Relationship experiments. The smallest datasets contained 50 subjects, but each dataset had at least 19 channels per subject amounting to a lower bound of 950 distinct subject-channels each with 40 features. From this lower subject-channel bound there the number of epochs in the training and enrollment datastets would change based upon the epoch duration. With the largest epoch duration of 10 seconds, there would be at minimum 9 epochs for each of the 950 subject-channels producing 8,550 unique subject-channel-epochs to model. This value exceeded the upper limits of the two mixture sweeps ensuring overfitting was not a major influence on performance.

3.3.2 TVM Dimensions

The size of the TVM was bounded by the number of mixtures and a depth factor called l from section **??**. As stated, this l value had to be less than or equal to the number of subjects, otherwise models would be built specifically for each subject. Overfitting concerns were addressed with respect to the mixture sizes, but limiting the TVM depth to the number of subjects assured overfitting was impossible in the production of I-Vectors.

This was not strictly required, as the examples used to inform this work would build TVMs with a depth beyond that of the number of subjects [28]. As the TVM is an intermediate step before finalizing the I-Vectors with Linear Discriminate Analysis (LDA), the dimension of the TVM could be 1200 for processing data from 75 subjects

[29]. However, such options were based on datasets with an order of magnitude more epochs and contained feature vectors double in size than was proposed in this work

Bounding the upper limit was necessary given the dynamic between mixture size, TVM depth, and LDA depth. An upper bound of 200 was chosen because the majority of datasets and aggregated datasets would not exceed 200 subjects. Additionally, producing the TVM was the most computational intense components of the algorithm requiring a tradeoff of the sweep range and execution time. The lower bound was set at 25, half the smallest subject count. Three incremental values were used to step between the lower and upper bound which resulted in the following sweep range: {25, 50, 75, 100, 200}.

3.3.3 LDA Dimension

The use of LDA to finalize the I-Vectors was well documented by the founders of I-Vectors [30, 28] highlighting their own sweep for optimization with speech data. Thus LDA depth represented a third parameter to consider when building and evaluating I-Vector performance. The upper bound of LDA is determined by the size of its paired TVM. As the range of TVM dimensions was being aligned with the various aggregated dataset subject counts, the LDA dimensions were aligned to operate on a similar scaling.

The lower bound for LDA size was set to 15, slightly less than the TVM lower bound, and the upper bound was set to 100, half the TVM upper bound. Five intermediate values were chosen between the bounds which resulted in the following sweep range: {15, 30, 45, 60, 75, 100}. By focusing on smaller increments this parameter was designed to be less influential than the mixture size and the TVM depth. This sweep range would later be adjusted following the results of the Parameter Sweeps to: {5, 15, 20, 25, 25, 50, 75, 95, 100, 150, 195}.

3.3.4 Epoch Configuration

The final controllable parameters were the number and duration of epochs. Drawing the experiments from the work of La Rocca et al, the initial epoch duration was 10 seconds with 6 epochs per subject, based around the resting trials of the PhysioNet Database. The epoch durations were expanded to include 5, 2, and 1 second epochs. This naturally altered the number of epochs as the PhysioNet Database contained 1 minute and 2 minute trials which split into a various numbers of epochs for each epoch duration recording combination, show in Table 3.4.

Table 3.4. Epoch Duration Configuration. Number of total epochs per subject as a function of epoch duration.

	Epoch Duration (s)			
Trial Duration (s)	10	5	2	1
60	6	12	30	60
120	12	24	60	120

Based upon reviewer feedback to a prior publication [31] epoch generation was altered to enabled the number of epochs to be independent of epoch duration. This provided another parameter to sweep, number of epochs, which was previously conflated with the epoch duration and trial duration.

3.3.5 Dataset-Feature

Every experiment conducted used all three feature sets, but not every combination of datasets was explored. This was because finding an optimal feature set was beyond the scope of the proposed work. There were not enough available resources in terms of datasets, features, and time to satisfy a robust feature search. However, it was understood that the proposed experiments could offer insight into feature selection which is why every experiment used all three feature sets.

Despite this limitation, using all three feature sets for each experiment provided a comparison point for understanding algorithm-dataset performance. It was hypothesized that one feature set would generally outperform the others, independent of data. Variations in relative performance triggered by mixture size, TVM depth, LDA depth, or epoch settings were used to define areas of interest with respect to these controllable parameters. Additionally, using multiple features mitigated any potential bias generated for stumbling upon an ideal dataset-feature-algorithm combination and being able to identify it as such given the number of dataset-feature-algorithm pairings.

3.4 Implementation

In keeping with the theme of publicly available datasets, the software and hardware solutions were developed to be open sourced. As the research intersects multiple communities it was important that access be given to all regardless of expertise in software development or hardware support. Many of the latest data science solutions required every updating tool kits running on large computing clusters which can limit the use of novel tool kits.

3.4.1 Software

The initial search for I-Vector toolboxes yielded bob.spear[32], Kaldi[33], and Microsoft Research (MSR) Identity Toolbox[34]. The bob.spear toolbox did not work on Windows based machines and Kaldi had proven difficult to implement on the NEDC computing cluster. However, the MSR Identity Toolbox was developed with MATLAB and was easily setup locally and on the computing cluster.

The majority of software was developed specifically for this research with minor components drawn from public sources. A MATLAB toolbox called VOICEBOX[3] was used to support handling of the CEP features generated as Hidden Markov Toolkit (HTK) files. All European Data Format (EDF) EEG files were manipulated using edfREAD available through Mathworks MATLAB File Exchange[4].

The decision was made to build using Matlab because it provided a known functional model in the MSR toolbox, would be accessible to both the speech processing and EEG communities, and be robust to hardware/software configurations, and scalable for use on computing clusters. In hindsight there were tradeoffs in terms of performance and flexibility that may have been mitigated by developing the software tools in Python, but the development of this software package was a tertiary goal. Over the duration of the research the Matlab versions started with R2015A and finished on R2017B.

A review of the major facets of the software's workflow is provided in this section. All experiments started with feature creation as the data already existed within the NEDC file system. Once features were produced, a parameter file was written to control the experiment. This file outlined how each process would operate. The experiments were run sequentially to assure each algorithm used the same randomly generated epoch splits for training, enrollment, and testing data. Ultimately, all experiments were run on the NEDC clustering requiring Bash scripts to interface

[3]`http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html`

[4]`https://www.mathworks.com/matlabcentral/fileexchange/31900-edfread`

with our Slurm Workload Manager. Those interested in the individual classes and functions should refer to public Git repository's ReadMe[5].

3.4.1.1 Feature Creation

The conversion of EEG recordings, stored as EDF files, into CEP, COH, and PSD features was independent of the experiments. This was done to ensure static feature sets and simplified the structure of processing the features during the experiments. Given the number of 'channels' produced from COH features, all feature data was indexed and saved in relation to their epochs.

Thus the number of files produced for each feature set was dependent on subject and number of epochs with channel data organized inside each epoch file. These file lists were the inputs to the experiments where they were aggregated. This tool was written to run with multiple Matlab workers and was supported via a Slurm base script.

3.4.1.2 UBM Class

The use of UBMs was handled through the development of a Universial Background Model (UBM) class in Matlab called *UniBacMod.m.* This simplified the generation, evaluation, and loading/saving of existing models. The generation of the UBMs leveraged Matlab's Single Program Multiple Data (SPMD) parallel computing feature to carry out the expectation maximiation (EM) process on the training data. The enrollment models were built using a parallel maximum a priori (MAP) adaptation from the generated models and enrollment data. These enrollment models were compared against the testing data to produce log-likelihood ratios which were scored for CRR and EER.

[5] `https://github.com/izlandman`

This class controlled the number of UBMs mixtures, the number of EM iterations, and the downsample factor. In addition it held the number of epochs and the resultant UBMs. All of these variables were saved after converting the class to a structure enabling subsequent I-Vector experiments to use the same UBMs.

3.4.1.3 TVM Class

The use of TVMs was handled through the development of a total variability matrix (TVM) class in Matlab called *TotVarMat.m*. This simplified the generation, evaluation, and loading/saving of existing models. Again the EM process used to build the TVM was run using the same parallel processes for the UBM class. The generation of I-Vectors was done in parallel as well, with the option to produce a set of LDA constrained I-Vector in addition to the native TVM I-Vectors. Final evaluations between the I-Vectors were carried out through a parallel cosine distance function to produce the CRR and EER metrics.

The class retained the enrollment and testing I-Vectors and performance metrics binary files, with all other parameters saved as a Matlab structure. Control over the depth of the TVM, depths of the LDA variants, and training steps for the TVM EM were carried out in this class. The constraints previously laid out by the imported UBM are inherited by the TVM class. This assured the number of epochs and UBM parameters were consistent between algorithms. Critically, this allowed the production of I-Vectors from a static UBM produced in a prior experiment.

3.4.1.4 Mahalanobis Evaluation

The use of Mahalanobis Distance as a classifier was borrowed from the work of La Rocca et al. [13]. They developed their experiments using Matlab using the built-in `Mahal` function from the Statistics and Machine Learning Toolbox. Each training/enrollment subject's epochs were used to produce a subject mean. The variances for each feature were drawn from a pooled covariance matrix built from all subject's epoch data.

Evaluation of the the distance matrix between all subjects was used to produce CRRs and EERs aligned to the same epoch, mixture, LDA depth process as the I-Vectors. The resultant distance matrices were saved for each step of the cross-validation process as a binary file. No class was built for this process as it was not the main focus of the proposed research.

3.4.2 Hardware

All of the experiments were run on the NEDC computing cluster, Neuronix. While the cluster supported CPU and GPU parallel processes, the toolkit was written to only support CPU parallelization. Neuronix contained four main identical CPU compute nodes and two minor identical CPU compute nodes. The main nodes consisted of two AMD Opteron 6378s with 16 cores supported by 128GB of DDR3 Ram. The minor nodes consisted of two Intel Xeon E5-2603s with 8 cores supported by 128GB of Ram.

The data server consisted of over 2TB of disk space shared by all the users of NeuroNix.

Chapter 4

NEAR FIELD COMMUNICATION BASED ACCESS CONTROL FOR WIRELESS MEDICAL DEVICES

Chapter 5

A PATIENT ACCESS PATTERN BASED ACCESS CONTROL SCHEME

Chapter 6

PATIENT INFUSION PATTERN BASED ACCESS CONTROL SCHEMES FOR WIRELESS INSULIN PUMP SYSTEM

Chapter 7

BIOMETRICS BASED TWO-LEVEL SECURE ACCESS CONTROL FOR IMPLANTABLE MEDICAL DEVICES DURING EMERGENCIES

Chapter 8

CONCLUSION

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals." *Circulation*, vol. 101, no. 23, pp. E215–20, jun 2000.

[2] I. Obeid and J. Picone, "The Temple University Hospital EEG Data Corpus." *Front. Neurosci.*, vol. 10, no. MAY, p. 196, may 2016.

[3] G. Schalk *et al.*, "BCI2000: a general-purpose brain-computer interface (BCI) system." *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–43, jun 2004.

[4] M. Delpozo-Banos, C. M. Travieso, C. T. Weidemann, and J. B. Alonso, "EEG biometric identification: A thorough exploration of the time-frequency domain," *J. Neural Eng.*, vol. 12, no. 5, 2015.

[5] S. Yang, F. Deravi, and S. Hoque, "Task sensitivity in EEG biometric recognition," *Pattern Anal. Appl.*, pp. 1–13, 2016.

[6] B. Reuderink, J. Farquhar, M. Poel, and A. Nijholt, "A subject-independent brain-computer interface based on smoothed, second-order baselining," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 4600–4604, 2011.

[7] M. Fraschini *et al.*, "An EEG-based biometric system using eigenvector centrality in resting state brain networks," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 666–670, 2015.

[8] D. Rodrigues *et al.*, "EEG-based person identification through Binary Flower Pollination Algorithm," *Expert Syst. Appl.*, vol. 62, pp. 81–90, nov 2016.

[9] K. Su and K. A. Robbins, "A Framework for Content-based Retrieval of EEG with Applications to Neuroscience and Beyond." *Proc. ... Int. Jt. Conf. Neural Networks. Int. Jt. Conf. Neural Networks*, pp. 1–8, 2013.

[10] F. Lotte *et al.*, "A review of classification algorithms for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 4, no. 2, pp. R1–R13, jun 2007.

[11] A. M. Dymond, R. W. Coger, and E. A. Serafetinides, "Preprocessing by factor analysis of centro-occipital EEG power and asymmetry from three subject groups," *Ann. Biomed. Eng.*, vol. 6, no. 2, pp. 108–116, 1978.

[12] C. Berka *et al.*, "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks." *Aviat. Space. Environ. Med.*, vol. 78, no. 5 Suppl, pp. B231–44, may 2007.

[13] D. La Rocca *et al.*, "Human brain distinctiveness based on EEG spectral coherence connectivity," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 9, pp. 2406–2412, 2014.

[14] M. V. Ruiz-blondet, Z. Jin, and S. Laszlo, "CEREBRE: A Novel Method for Very High Accuracy Event-Related Potential Biometric Identification," *IEEE Trans. Inf. Forensics Secur.*, vol. 6013, no. c, pp. 1–13, jul 2016.

[15] S. Furui, "Cepstral Analysis Technique for Automatic Speaker Verification," *IEEE Trans. Acoust.*, vol. 29, no. 2, pp. 254–272, 1981.

[16] H. Li, B. Ma, and K.-A. Lee, "Spoken Language Recognition: From Fundamentals to Practice," *Proc. IEEE*, vol. 101, no. 5, pp. 1136–1159, may 2013.

[17] A. Harati *et al.*, "Improved EEG event classification using differential energy," in *2015 IEEE Signal Process. Med. Biol. Symp. - Proc.*, no. December 2015. IEEE, dec 2016, pp. 1–4.

[18] S. Lopez *et al.*, "Automated identification of abnormal adult EEGs," in *2015 IEEE Signal Process. Med. Biol. Symp.*, vol. 37, no. 6. IEEE, dec 2015, pp. 1–5.

[19] V. J. Monastra, J. F. Lubar, and M. Linden, "The development of a quantitative electroencephalographic scanning process for attention deficit–hyperactivity disorder: Reliability and validity studies." *Neuropsychology*, vol. 15, no. 1, pp. 136–144, 2001.

[20] S. Makeig *et al.*, "Evolving signal processing for brain-computer interfaces," in *Proc. IEEE*, vol. 100, no. SPL CONTENT, aug 2012, pp. 1567–1584.

[21] B. Güntekin, E. Ba, and E. Ba←(s) ar, "Review of evoked and event-related delta responses in the human brain," *Int. J. Psychophysiol.*, vol. 103, pp. 43–52, 2016.

[22] S. Marcel, J. D. R. Millán, and J. d. R. Millan, "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 743–748, 2007.

[23] P. Nguyen *et al.*, "EEG-Based person verification using Multi-Sphere SVDD and UBM," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 7818 LNAI, no. PART 1, pp. 289–300, 2013.

[24] M. H. Bahari, M. McLaren, H. Van Hamme, and D. A. Van Leeuwen, "Speaker age estimation using i-vectors," in *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 1, sep 2012, pp. 506–509.

[25] O. Glembek *et al.*, "Simplification and optimization of i-vector extraction," in *2011 IEEE Int. Conf. Acoust. Speech Signal Process.* IEEE, may 2011, pp. 4516–4519.

[26] H. Behravan, V. Hautamäki, and T. Kinnunen, "Foreign Accent Detection from Spoken Finnish Using i-Vectors," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, no. August, pp. 79–83, 2013.

[27] J. L. Marcano, M. A. Bell, and A. L. Beex, "Classification of ADHD and non-ADHD subjects using a universal background model," *Biomed. Signal Process. Control*, vol. 39, pp. 204–212, 2018.

[28] N. Dehak *et al.*, "Front-End Factor Analysis for Speaker Verification," *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 19, no. 4, pp. 788–798, may 2011.

[29] M. McLaren and D. van Leeuwen, "Improved speaker recognition when using i-vectors from multiple speech sources," *2011 IEEE Int. Conf. Acoust. Speech Signal Process.*, no. 1, pp. 5460–5463, may 2011.

[30] N. Dehak, P. A. Torres-Carrasquillo, D. A. Reynolds, and R. Dehak, "Language Recognition via i-vectors and Dimensionality Reduction," in *INTERSPEECH*, no. August, 2011, pp. 857–860.

[31] C. Ward and I. Obeid, "Application of identity vectors for EEG classification," *J. Neurosci. Methods*, vol. 311, pp. 338–350, jan 2019.

[32] E. Khoury, L. E. Shafey, and S. Marcel, "Spear: An open source toolbox for speaker recognition based on Bob," in *2014 IEEE Int. Conf. Acoust. Speech Signal Process.* IEEE, may 2014, pp. 1655–1659.

[33] D. Povey *et al.*, "The Kaldi speech recognition toolkit," in *IEEE 2011 Work. Autom. speech Recognit. Underst.*, Cambridge, 2011.

[34] S. O. Sadjadi, M. Slaney, and L. Heck, "MSR identity toolbox v1. 0: a matlab toolbox for speaker-recognition research," *Speech Lang. Process. Tech. Comm. Newsl.*, no. 1, pp. 4–7, 2013.

APPENDIX A

AppendixA

A.1 Start Here

A.1.1 More Here

A.1.2 And Again

A.2 Restart!

APPENDIX B

Appendix2

B.1 Start Here

B.1.1 More Here

B.1.2 And Again

B.2 Restart!