Arrhythmia signatures with empirical mode decomposition

Ashwin Sundar

December 5, 2016

M.S. Applied Project
Chair: Dr. Jeffrey La Belle
Member: Dr. Mark Spano
Member: Dr. Heather Ross

Table of Contents

Acknowledgments

I'd like to thank Dr. Jeffrey La Belle, Dr. Mark Spano, and Dr. Heather Ross for their guidance and support. I'd also like to thank my family for their support in my education.

Abstract

Frequency domain techniques are often used to deconstruct and understand the electrocardiogram (ECG). However, ECG artifacts are brief, making frequency domain analysis challenging. An alternate signal processing technique, empirical mode decomposition (EMD), could be powerful because analysis stays in the time domain. EMD was applied to more than 1,500 waveforms from the MIT-BIH Arrhythmia Database, spanning a range of subjects and arrhythmia types. Physician annotations were used to window and sort waveforms, and EMD was used to deconstruct waveforms into intrinsic mode functions (IMFs). An average IMF for each type was calculated. IMFs were compared across arrhythmias types to identify unique signatures. These unique signatures could potentially be used to identify arrhythmias in the ECG. If capable, EMD for arrhythmia detection could be an attractive alternative to time-frequency analysis in portable, low-power systems.

Keywords: electrocardiography (ECG), empirical mode decomposition (EMD)

Specific Aims

Find signatures in ECG data that can be used to identify arrhythmias.

Introduction

History of electrocardiography

Italian physician and physicist Luigi Galvani first observed electrical activity in the late 18th century. His assistant accidentally touched an exposed sciatic nerve of a frog with a scalpel holding a static charge, and the frog's legs kicked, as if alive. This inspired Galvani to coin the term animal electricity to describe his belief that the fibers activated because of a certain type of electricity intrinsic to muscle tissue. One of Galvani's contemporaries, Alessandro Volta, argued that the contractions were due to an electrical effect of the metal wire itself (Bresadola, 1998).

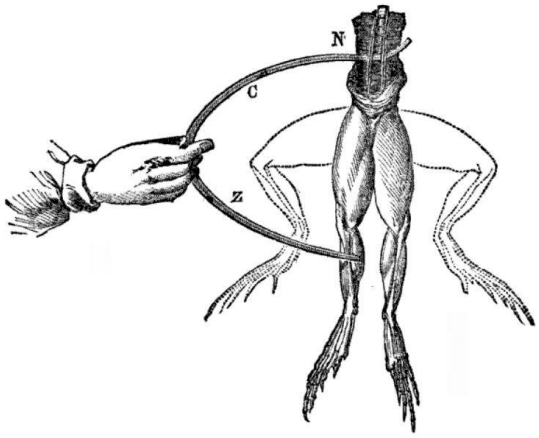


Figure 1: Electrodes contact frog muscles, causing the legs to twitch (Wells, 1859).

Inspired by the work of Galvani and Volta, physicist Carlo Matteucci further investigated animal electricity with his creation of a "frog battery", or a battery made from frog legs connected in series. During his work, Matteucci discovered that electrical signals accompanied each heart beat (Matteucci, 1842). British physiologists continued to characterize this electrical activity over the remainder of the century (AlGhatrif & Lindsay, 2012).

Modern electrocardiography began with Willem Einthoven's capillary electrometer, which could detect deflections in the electrical readings taken from the skin's surface. He called these deflections P, Q, R, S, and T, respectively, and coined the term "electrocardiogram" to describe this measurement technique (Einthoven, 1895). This, along with the development of Einthoven's triangle, earned him the 1924 Nobel Prize in Physiology or Medicine.

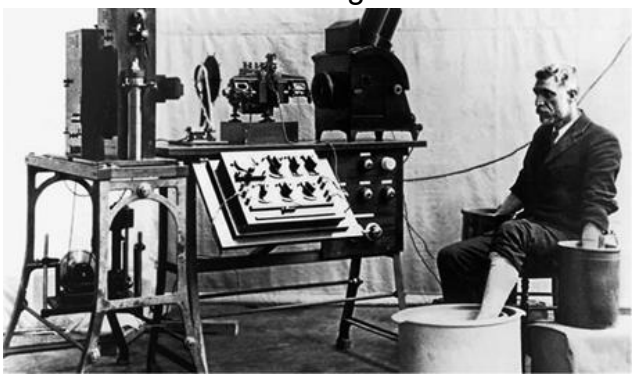


Figure 2: Einthoven's highly sensitive galvanometer was used to amplify the millivolt ECG signal, for analysis by his capillary electrometer. The full system weighed 600 pounds (Einthoven, 1895).

Further improvements in Einthoven's ECG were made over the next few decades and resulted in the development of unipolar leads and the 12-lead electrocardiogram as it is known today.

Physiology

The major feature of the heart that makes ECG possible is its well-characterized electrical system. This system, also known as the conduction system, controls the expansion and contraction of cardiac muscle. Features of the ECG correspond directly to these cardiac muscle movements. Many cardiac diseases are diagnosed based on the changes observed in the ECG.

In the healthy heart, an electrical signal originates in the sinoatrial (SA) node. This signal occurs independently of the brain, and is commonly referred to as the "natural pacemaker" of the heart (Campbell, Reece, Taylor, & Zalisco, 2006). From there, the signal propagates radially across the atria, causing the atria to contract. This contraction squeezes blood through the open valves and into the ventricles. Once the signal reaches the atrioventricular (AV) node near the ventricles, it briefly slows to allow the ventricles to fill with blood. From here, the signal travels to the Bundle of His, located within the ventricles. The signal then bifurcates into the left and right bundle branches through the septum, and then propagates across the ventricles via the Purkinje fibers, causing them to contract. The contraction of the right ventricle conveys deoxygenated blood to the lungs via the pulmonary valve, while the contraction of the left ventricle conveys oxygenated blood to the rest of the body through the aortic valve. Finally, the ventricles relax before the next signal (National Heart, Lung, and Blood Institute, 2011).

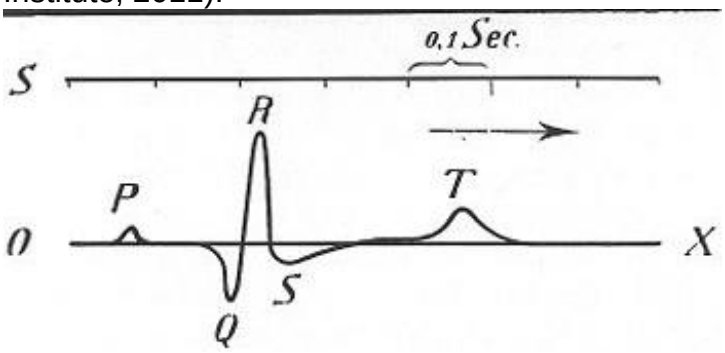


Figure 3: Einthoven likely chose the letters P, Q, R, S, and T to *follow Descartes' mathematical convention (Einthoven, 1895).*

As discussed by Einthoven, the major deflections of the healthy ECG are the P, Q, R, S, and T waves (Einthoven, 1895). Atrial contraction is marked by the P wave. The lull between the P and Q wave occurs because the signal is slowing to allow the ventricles to fill. Signal bifurcation is marked by the Q wave, and the contraction of the right ventricle is marked by the R wave. The contraction of the left ventricle is marked by the S wave, and finally the relaxation of the ventricles is marked by the T wave.

Math

Frequency domain analysis refers to a mathematical transformation that translates a set of data from the time domain to the frequency domain. The most common type of time-frequency transformation is the Fourier transform. The Fourier transform describes what combination of sine and cosine functions to add together to create the original signal.

Frequency domain analysis of ECG data is not new. A conservative search of "time-frequency domain" AND "ECG" on Google Scholar produces more than 5,000 hits in the last decade. Similarly, "Fourier transform" AND "ECG" results in more than 17,000 hits, and "wavelet transform" AND "ECG" results in more than 16,000 hits.

The technique explored in this paper is called empirical mode decomposition (EMD). It relies on the fact that a signal is equal to the sum of its parts. EMD pulls a signal apart into its components per an algebraic set of rules. This means that the signal never leaves the time domain.

$$S = \sum_{1}^{m} f_m(t) + R$$

Equation 1: The sum of the IMFs plus the residue equals the original signal. S = original signal; m = # of modes; R = residue, or error

EMD was developed as a joint effort by NASA, NOAA, Naval Research Laboratory, and several universities. The motive was to develop an adaptive and efficient technique to handle nonlinear and nonstationary data (Huang, et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, 1998). Features of ECG signals include baseline drift, power line noise, motion artifact, temporally short QRS complexes, and temporally short periods of cardiac abnormality. Fourier transforms struggle to accurately represent such messy data. EMD makes no assumptions about the data, and is always an accurate representation of the original signal.

Unlike the Fourier and wavelet transforms, EMD is more like an algorithm that can be applied to data sets, rather than a theoretical tool. The implementation used in this project was cubic-spline interpolation (CSI), as used in the seminal paper by Huang et al. (1998). Cubic spline curves are fit to the maxima and minima of the original signal, creating an envelope. The mean of the envelope is subtracted from the original signal, and the process is repeated until the standard deviation between the current and last signal falls below a threshold (0.3). This forms the first intrinsic mode function (IMF). This process is repeated until CSI can no longer be used to form an envelope. This final signal is called the residue (Huang, et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, 1998).

Some current biomedical applications of EMD include analysis of EEG response to transcranial magnetic stimulation (Pigorini, et al., 2011) and identification of seizure in EEG signals (Pachori, 2008). EMD has also seen applications in fields ranging from finance to seismology. Any field with real-world data could benefit from looking through the lens of EMD.

Methods

46 patient records from the MIT-BIH Arrhythmia Database were used for analysis. Records from patients with pacemakers were omitted. Each record contained data from two different lead configurations. Since MLII was the most common lead configuration, only MLII ECG data was used. A selection of physician annotations was identified within each record, and a 200-millisecond window was created around the annotation. This is because the QRS complex in a healthy patient typically spans 60 to 100 milliseconds (Sivaraks & Ratanamahatana, 2014). From each record, at least 10 normal sinus rhythm beats were captured.

Data needed to be parsed, and physician annotations had to be located manually. Data parsing was automated in MATLAB using the Waveform Database (WFDB) plugin, supplied by PhysioBank.

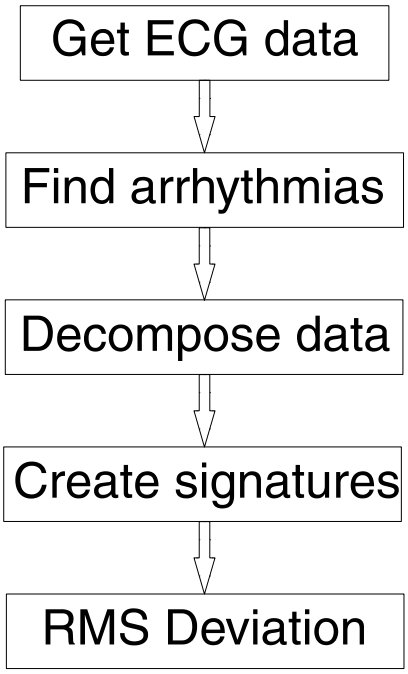


Figure 4: Flow of MATLAB code. First, ECG data is retrieved from PhysioBank using a MATLAB software package. Second, arrhythmias are located using physician annotations. Third, the ECG is decomposed into intrinsic mode functions, or IMFs. Fourth, the IMFs are averaged to create signature mode functions for each arrhythmia. Finally, the root mean square deviation is calculated to compare signatures from different arrhythmias.

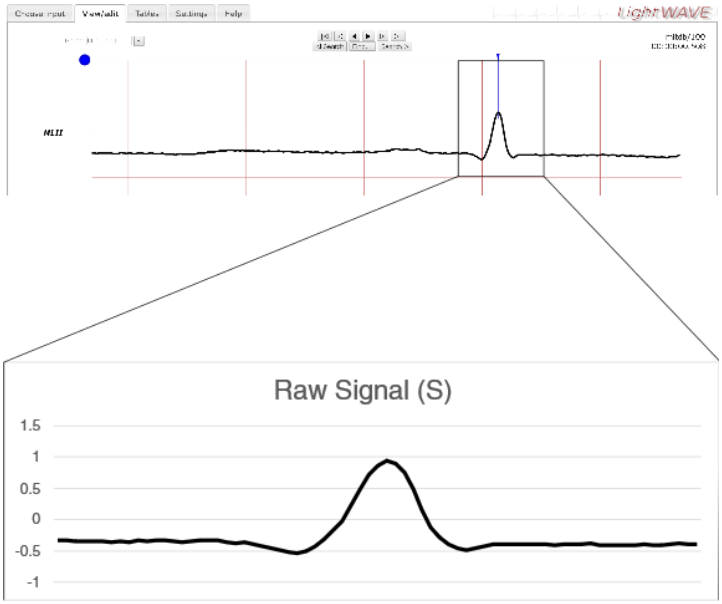


Figure 5: LightWAVE was used to located points of interest, such as arrhythmias. WFDB and Matlab were used to create a window around each point of interest.

Get ECG Data: First, patient ECG data was retrieved from PhysioBank using WFDB. The full patient record was stored in a 650000x2 matrix. The first column was time, and the second column was unitless raw data. MATLAB precision was set to 3 decimal places to take advantage of the precision of the incoming ECG data. Finally, the ECG from each patient was stored in its own .csv file.

Find arrhythmias: Next, arrhythmias were located using physician annotations in LightWAVE, an ECG viewing tool based in the web browser. Arrhythmia times in each record were noted in Excel, and MATLAB was used to retrieve arrhythmia windows based on those times. A 200-millisecond window was created around each arrhythmia.

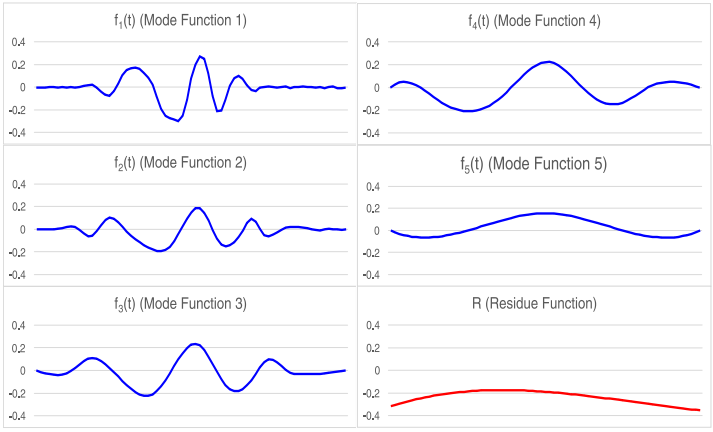


Figure 6: The raw signal in the above figure was decomposed into five IMFs and one residue function. Notice how early modes contain high frequency content, while later modes contain general trend information.

Decompose data: Next, signals were decomposed using a publicly-available MATLAB script (Rilling & Flandrin, 2007). First, the local extrema are located. Next, separate spline curves are fit to the maxima and minima, and the mean function of the envelope is calculated. Finally, this mean is subtracted from the starting signal to create an iterated signal. In the first iteration, the starting signal is the original signal itself. After the first iteration of this sifting process, the starting signal becomes the iterated signal from the last sift. This process of subtracting the mean continues, until the mean function has a standard deviation of less than 0.3. The resulting signal is the first intrinsic mode function, or IMF. The IMF is subtracted from the original signal, and the above process is repeated to generate the next IMF.

This entire procedure continues, and IMFs are generated until there are no longer enough local maxima or local minima to fit a spline curve. To summarize:

1. For a signal S(t), let m_1 be the mean of upper and lower envelopes of S(t). The envelopes are found using cubic-spline interpolation of the local maxima and minima.

2. The first iteration h_1 of the first IMF is computed:
$$h_1 = S(t) - m_1$$

3. In the second iteration, h_1 is now treated as the "data" set, and m_{11} is the upper and lower envelope mean:
$$h_{11} = h_1 - m_{11}$$

4. The process is repeated until the standard deviation of m_{1k} is less than 0.3

5. The next IMF is computed using the above process. IMFs are generated, until there are no longer enough maxima and minima to fit a spline curve.

Create signatures: Next, signatures were created. Several samples from each arrhythmia were decomposed and an average mode function, or signature, was created for each mode.

However, signals do not always decompose into the same number of modes, depending on the complexity. To address this issue, only the modes common to every signal were averaged. For example, take signal A and signal B. A decomposes into 5 modes, while B decomposes into 4 modes. In the averaging step, only the first 4 modes of A are averaged with the 4 modes of B.

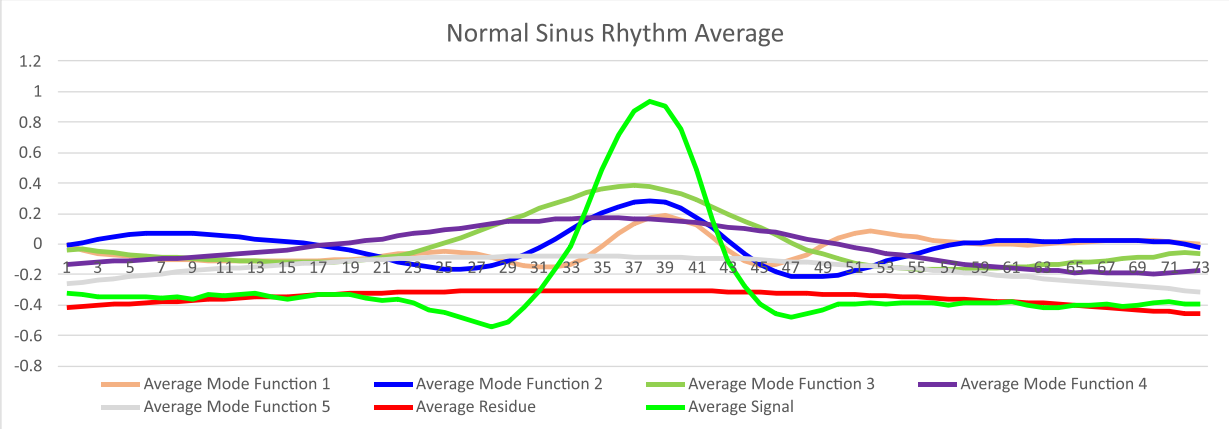


Figure 7: The average signal for normal sinus rhythm was calculated, based on 397 individual signals from 38 different patients. However, each signal does not always decompose into the same number of modes. Therefore, the averages for only the first five modes were calculated, since these were common to every signal.

$$A_{ak} = \frac{1}{n}\sum_{1}^{n} f_{nk}(t)$$

Equation 2: The signature modes for each arrhythmia are calculated by taking an average of each mode for all sample windows. k=mode number; A_{ak} = k^{th} average mode for arrhythmia a; n = # of sample windows.

Tables 1 and 2 in the Appendix contain more details and metadata for each patient record.

Root mean square (RMS) deviation: Finally, the signatures for each arrhythmia were compared by calculating the root mean square (RMS) deviation between each signature. This value was normalized by setting the means of each signal equal.

$$D_{abk} = sqrt(\sum_{1}^{73} \frac{1}{73}(f_{ak}(t) - f_{bk}(t) + (\overline{f_{ak}} - \overline{f_{bk}}))^2)$$

Equation 3: To compare the mode functions between arrhythmias, the root mean square deviation was calculated. The means were set equal and the point-by-point distance was calculated. The sum of the square root of each distance results in the RMS deviation. D_{abk}: RMS deviation of mode k of arrhythmia a versus mode k arrhythmia b; f_{ak}: mode function k of arrhythmia a; f_{bk}: mode function k of arrhythmia b; $\overline{f_{ak}}$: mean of mode function k of arrhythmia a; $\overline{f_{bk}}$: mean of mode function k of arrhythmia

Results

To visualize the results, the root mean square deviations between signatures were placed in a matrix. Each cell contains the RMS deviation between the arrhythmia type in the top row and the arrhythmia type in the left column. The full matrix containing all 20 arrhythmia types is listed in Figures 10, 11, and 12 in the Appendix.

To help focus the comparison, the signatures derived from at least 10 patients and 100 sample window is shown in Figure 8.

Ventricular bigeminy drew from 171 data sets and 13 patients. Atrial premature beat drew from 319 data sets and 27 patients. Fusion of ventricular and normal beat drew from 136 data sets and 17 patients. Normal sinus rhythm 397 data sets and 38 patients. Finally, premature ventricular contraction drew from 593 data sets and 36 patients.

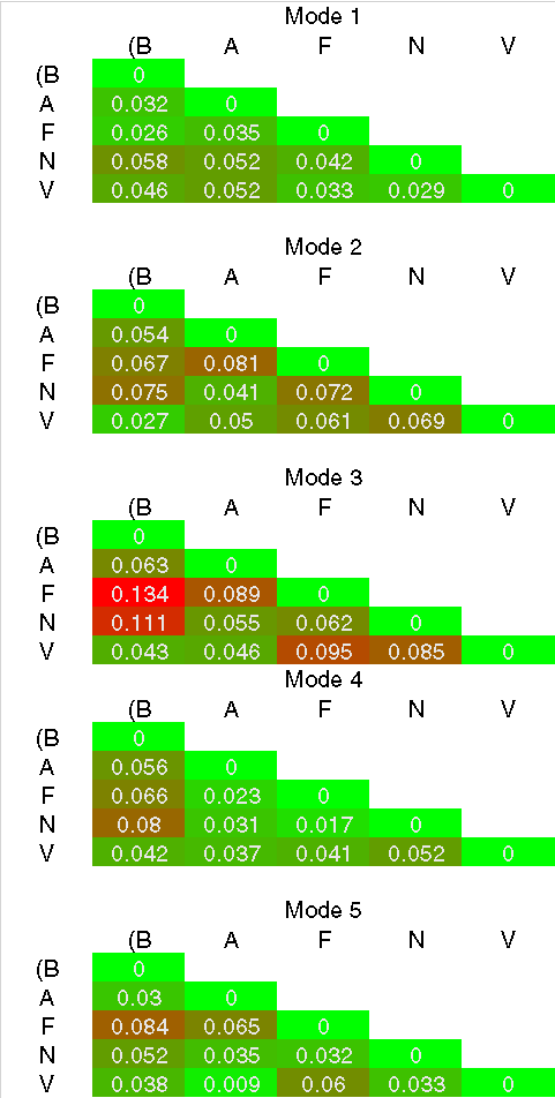


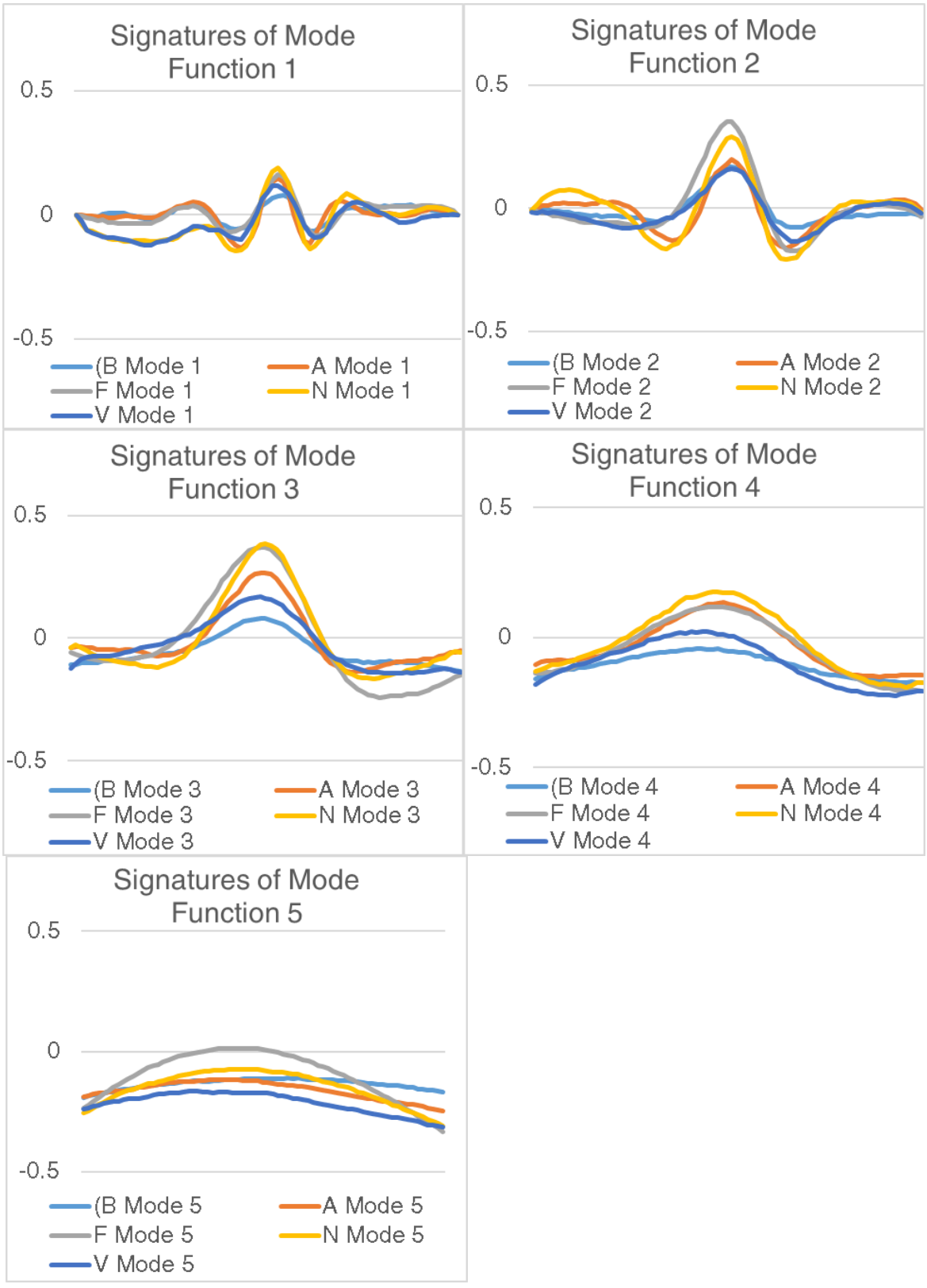
Figure 8: The RMS deviations between four arrhythmia types and normal sinus rhythm is shown here. The diagonal always equals 0, and the upper diagonal is equal to the lower diagonal. Abbreviations are:
(B: Ventricular bigeminy
A: Atrial premature beat
F: Fusion of ventricular and normal beat
N: Normal sinus rhythm
V: Premature ventricular contraction

As expected, comparing a signature to itself should result in an RMS deviation of 0, which is shown along the diagonal of each matrix. Additionally, the upper diagonal of the matrix is equal to the lower diagonal.

The matrices are color coded from green to red. Green indicates the highest degree of similarity, which is 0 in each matrix. Red indicates the highest degree of difference, which is 0.134 in Figure 8. Mode function 3 appeared to show a relatively wide range of differences. This includes the largest observed RMS deviation, between ventricular bigeminy and fusion of ventricular and normal beat.

In Figure 9, each set of signatures is plotted as well. The earlier modes feature high-frequency content, while the later modes feature more low-frequency content.

Figure 9: The signatures from ventricular bigeminy, atrial premature beat, fusion of ventricular and normal beat, normal sinus rhythm, and premature ventricular contraction are plotted here.



Discussion

In Figure 8, it appears that there exist the greatest similarities between arrhythmia types in mode 1, while mode 3 appears to bear the greatest differences. Additionally, modes 1 and 2 appear to contain more noise content, while modes 3-5 appear to contain information on the general trend of the data. Perhaps this means that the trend of each arrhythmia is most prominent in mode 3. To shed more light on this idea, other arrhythmia types should be analyzed as well and added to the chart. It could be the case that among these arrhythmias, the signature differences are most apparent in mode 3.

In Figures 10, 11, and 12 in the Appendix, it appears as if a few arrhythmia types are extremely different from all other arrhythmia types. Supraventricular premature beat and 2° heart block appear to diverge significantly. On closer inspection, Table 1 in the Appendix indicates only 2 sample windows from 1 patient contributed to supraventricular premature beat, while only 5 records from 1 patient contributed to 2° heart block. Therefore, these signatures likely appear different because of the low sample size.

Additionally, in Figures 10, 11, and 12 in the Appendix, it appears that mode 1 and mode 5 bear the greatest similarities between arrhythmia types. This could mean that high-frequency content and basic trend information in all arrhythmia types tends to be similar. Once again, mode 3 shows the greatest differences between arrhythmia types. This further contributes to the idea that relevant arrhythmia signatures could possibly be drawn from this mode.

Unlike the Fourier transform, empirical mode decomposition is not a theoretical technique. Rather, it is an observation that a signal can be broken down into the sum of its parts. Cubic-spline interpolation is just one method for breaking down a signal. Other proposed methods include moving least-squares approximation (Blakely, 2005) and PDE-based sifting (Deléchelle, Lemoine, & Niang, 2005).

There are some potential improvements that could be made to this process. Signatures were created by calculating the average of the modes of the sample windows. The signals were normalized by setting the means equal. This step could be made more rigorous by also normalizing the ranges of the data as well.

Another possible improvement could be made by changing the length of the windows. All windows were statically set to be 100 milliseconds long, since the average QRS complex spans 100 milliseconds (Sivaraks & Ratanamahatana, 2014). However, this number relies on the length of the healthy QRS complex, and in many arrhythmias the QRS complex has a different shape and spans a different amount of time. A survey of the literature for each arrhythmia would be needed to make a more informed decision on how long windows should be for every arrhythmia type.

A set of data may be ill-conditioned for Fourier transformation if it is non-linear, non-stationary, and unevenly spaced. Discrete Fourier transform algorithms are lossy – computer rounding error and finite sampling rates compound this issue. Regardless, most data analysis in engineering continues to use the Fourier transform because frequency information tends to be valuable.

EMD imposes no limitations on incoming data, which is important in the world of electrocardiology. Real world ECG data is noisy, drifting, and unevenly spaced. Slight changes in lead placement can result in drastic changes in the ECG. Anatomy and skin

impedance can be highly variable between patients and under different conditions. These are just a few reasons why arrhythmia-detection algorithms are difficult to create.

Conclusion

The analysis described above is just a suggestion on what intrinsic mode functions potentially mean. The goal was to create a training data set of signatures that could be used to identify arrhythmias in unknown electrocardiograms, independent of an electrocardiologist. One way to achieve this goal is to decompose an annotated signal, and see how well its IMFs match to the signatures described above.

If capable, EMD for arrhythmia detection could be an attractive alternative to time-frequency analysis in portable, low-power systems. Advancements in computing in the last decade have expanded the availability of cheap and easy-to-use microcontroller platforms, such as Arduino Uno. Time-frequency analysis is powerful, but computationally expensive on most of these platforms. Computational complexity consumes battery power, which is the biggest limitation of any portable microcontroller platform.

Empirical mode decomposition is a relatively new algorithmic technique for processing data. The major benefit is that EMD is intuitive to understand – the original signal is the sum of its intrinsic mode functions and the residual. Every mode contains information about the state of the signal at every given time. Additionally, EMD makes no assumptions about data – it can work for any set of data, without requirements of linearity or stationarity. Finally, EMD can be relatively easy to implement in a microcontroller environment, and could making remote arrhythmia detection a reality.

Learning Objectives

Data from the U.S. healthcare system has reached the hundreds of exabytes, equal to the several thousand copies of every item in Library of Congress (Wullianallur & Raghupathi, 2014). I think the human condition can be greatly improved by finding better ways to analyze and understand all this data. The past decade has seen the explosion of the smartphone, an example of a cheap, portable, and powerful computer. That said, data processing does not need to have a powerful computer behind it, as I began to demonstrate in this paper.

PhysioBank has made anonymized patient health data publicly available for decades. I wanted to find a new way to process the data. A quickly literature search revealed that time-frequency analysis of physiological data is extremely popular. A lengthy literature search suggested that empirical mode decomposition of arrhythmias in ECG data was not as common. I decided this may be a good place to begin. Along the way, I wished to improve my programming skills, extend my knowledge of physiological data processing, and learn how to come up with solutions to problems that don’t exist in textbooks.

References

AlGhatrif, M., & Lindsay, J. (2012). A brief review: history to understand fundamentals of electrocardiography. *Journal of Community Hospital Internal Medicine Perspectives*.

Blakely, D. C. (2005). A Fast Empirical Mode Decomposition Technique for Nonstationary Nonlinear Time Series. *Elsevier Science*.

Bresadola, M. (1998). Medicine and science in the life of Luigi Galvani (1737–1798). *Brain Research Bulletin*, 367-380.

Campbell, N. A., Reece, J. B., Taylor, M. R., & Zalisco, E. J. (2006). *Biology: Concepts & Connections*. Pearson.

Deléchelle, E., Lemoine, J., & Niang, O. (2005). Empirical Mode Decomposition: An Analytical Approach for Sifting Process . *IEEE Signal Processing Letters*, 764-767.

Einthoven, W. (1895). Ueber die Form des menschlichen electrocardiogramms. *Pflügers Archiv: European Journal of Physiology*, 101-123.

Huang, N. E., Shen, Z., & Long, S. R. (1999). A new view on nonlinear water waves: The Hilbert Spectrum. *Annual Review of Fluid Mechanics*, 417-457.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., . . . Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society*, 903-995.

Matteucci, C. (1842). Sur un phenomene physiologique produit par les muscles en contraction. *Annales de chimie et de physique*, 339-341.

National Heart, Lung, and Blood Institute. (2011, November 17). *Your Heart's Electrical System*. Retrieved from http://www.nhlbi.nih.gov/health/health-topics/topics/hhw/electrical

Pachori, R. B. (2008). Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. *Research Letters in Signal Processing*, 1-5.

Pigorini, A., Casali, A. G., Casarotto, S., Ferrarelli, F., Baselli, G., Mariotti, M., . . . Rosanova, M. (2011). Time-frequency spectral analysis of TMS-evoked EEG oscillations by means of Hilbert-Huang transform. *J Neurosci Methods*, 236-245.

Rilling, G., & Flandrin, P. (2007). *Empirical Mode Decomposition*. Retrieved from http://perso.ens-lyon.fr/patrick.flandrin/emd.html

Sivaraks, H., & Ratanamahatana, C. A. (2014). Robust and Accurate Anomaly Detection in ECG Artifacts Using Time Series Motif Discovery. *Computational and Mathematical Methods in Medicine*.

Wells, D. A. (1859). *The science of common things: a familiar explanation of the first principles of physical science. For schools, families, and young students*. Ivison, Phinney, Blakeman.

Wullianallur, R., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*.

Appendix

Symbol	Arrhythmia	# of patient records	# of data sets
!	Ventricular flutter wave	1	42
(AFIB	Atrial fibrillation	7	78
(AFL	Atrial flutter	3	36
(B	Ventricular bigeminy	13	171
(BII	2° heart block	1	5
(IVR	Idioventricular rhythm	2	4
(NOD	Nodal (A-V junctional) rhythm	3	7
(SVTA	Supraventricular tachyarrhythmia	7	26
(T	Ventricular trigeminy	12	77
(VFL	Ventricular flutter	1	6
(VT	Ventricular tachycardia	12	54
A	Atrial premature beat	27	319
E	Ventricular escape beat	2	30
F	Fusion of ventricular and normal beat	17	136
L	Left bundle branch block beat	4	82
N	Normal sinus rhythm	38	397
R	Right bundle branch block beat	6	132
S	Supraventricular premature beat	1	2
V	Premature ventricular contraction	36	593
x	Non-conducted P-wave (blocked APB)	5	86

Table 1: Arrhythmias, abbreviations, and number of samples.

Record	Lead configuration	Medications	Points of Interest	Normal Sinus Rhythm?
100	MLII	Aldomet | Inderal	V | A	Y
101	MLII	Diapres	A	Y
103	MLII	Diapres | Xyloprim	V | (NOD	Y
105	MLII	Digoxin | Nitropaste | Pronestyl	V	Y
106	MLII	Inderal	V | (B | (T | (VT	Y
107	MLII	Digoxin	V	N
108	MLII	Digoxin | Quinaglute	A | V | F | x	Y
109	MLII	Quinidine	L | V | F	N
111	MLII	Digoxin | Lasix	L | V	N
112	MLII	Digoxin | Pronestyl	A	Y
113	MLII		a	Y
114	MLII	Digoxin	A | J | V | F | (SVTA	Y
115	MLII			Y
116	MLII		A | V	Y
117	MLII		A	Y
118	MLII	Digoxin | Norpace	R | A | V | x	N
119	MLII	Pronestyl	V | (B | (T	Y
121	MLII	Digoxin | Isordil | Nitropaste	A | V	Y
122	MLII	Digoxin | Lasix | Pronestyl		Y
123	MLII	Digoxin | Inderal	V	Y
124	MLII	Digoxin | Isordil | Quindine	R | A | J | V | F | j | (IVR | (NOD | (T	N
200	MLII	Digoxin | Quinidine	A | V | F | (B	Y
201	MLII	Digoxin | Hydrochlorthiazide | Inderal | KCl	A | a | J | V | F | j | x | (AFIB | (NOD | (SVTA | (T	Y
202	MLII	Digoxin | Hydrochlorthiazide | Inderal | KCl	A | a | V | F | (AFIB | (AFL	Y
203	MLII	Coumadin | Digoxin | Heparin | Hygroton | Lasix	a | V | F | (AFL | (T | (VT	Y
205	MLII	Digoxin | Quinaglute	A | V | F | (VT	Y
207	MLII	Digoxin | Quinaglute	L | R | A | V | ! | E | (B | (IVR | (SVTA | (VFL | (VT	N
208	MLII		S | V | F | (T	Y
209	MLII	Aldomet | Hydrodiuril | Inderal	A | V | (SVTA	Y
210	MLII		a | V | F | E | (AFIB | (B | (T | (VT	Y
212	MLII		R	Y
213	MLII	Digoxin	A | a | V | F | (B | (VT	Y
214	MLII	Digoxin | Dilantin	L | V | F | (T | (VT	N
215	MLII		A | V | (VT	Y
217	MLII	Digoxin | Lasix | Quinidine	V | F | (AFIB | (B | (VT	Y
219	MLII	Digoxin	A | V | F | x | (AFIB | (B | (T	Y
220	MLII	Digoxin	A | (SVTA	Y
221	MLII	Hydrochlorthiazide | Lasix	V | (AFIB | (B | (T | (VT	Y
222	MLII	Digoxin | Quinidine	A | J | j | (AB | (AFIB | (AFL | (NOD | (SVTA	Y
223	MLII		A | a | V | F | e | (B | (T | (VT	Y
228	MLII	Digoxin | Norpace	A | V | (B	Y
230	MLII	Dilantin	A | V | (B	Y
231	MLII		A | V | x | R | (BII	Y
232	MLII	Aldomet | Inderal	R | A | j	N
233	MLII	Dilantin	A | V | F | (B | (T | (VT	Y
234	MLII		J | V | (SVTA	Y

Table 2: Patient record metadata. Patient records contained data from two lead configurations. MLII was chosen since it was the most common lead configuration. Different lead configurations will result in a different ECG pattern. The abbreviations for the points of interest are given in Table 1. The final column indicates whether the record had any instances of normal sinus rhythm.

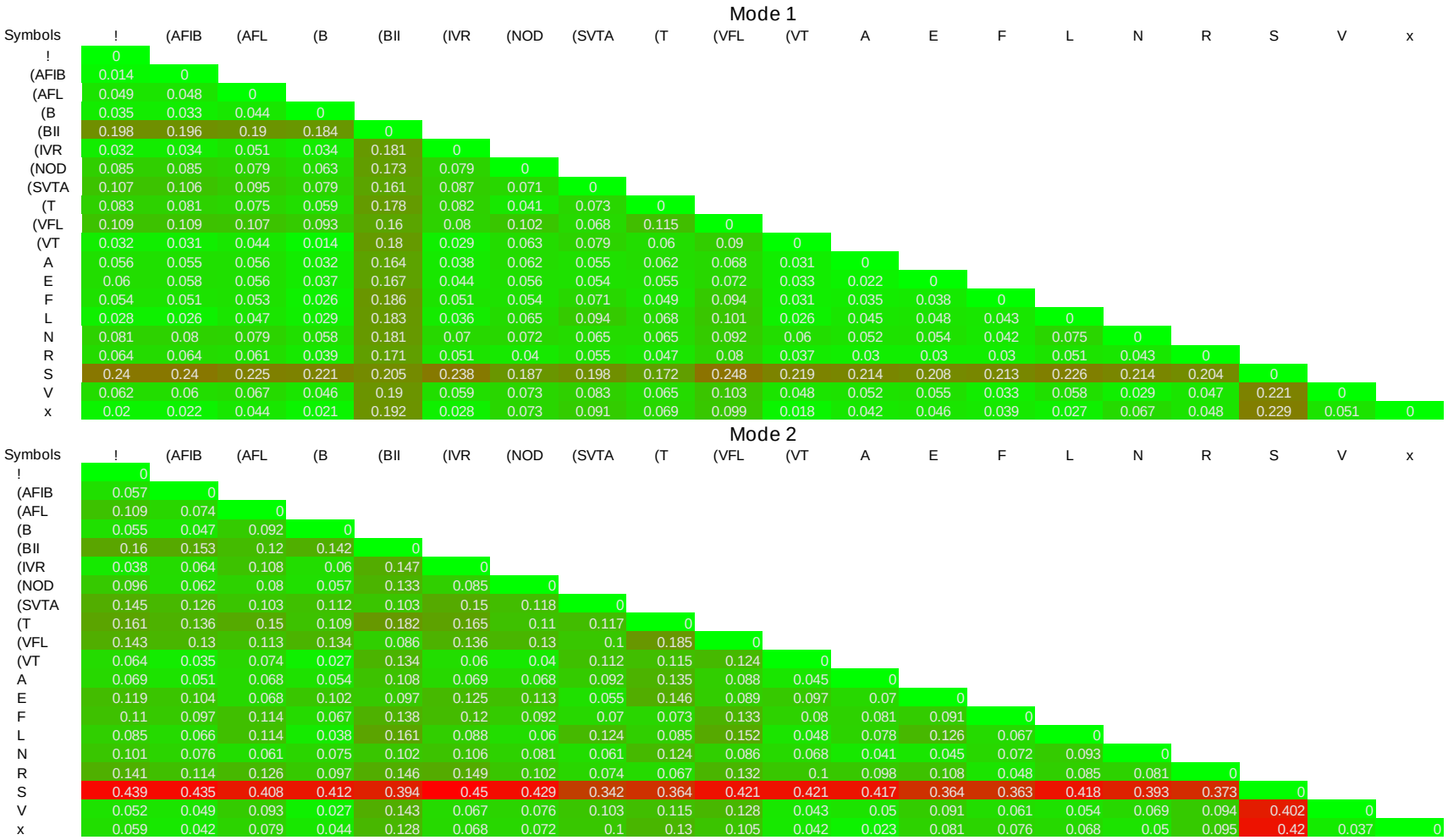


Figure 10: The differences between mode 1 for every arrhythmia type are captured in the top matrix. These values are the RMS deviations. Green indicates an exact match between the signal, while red represents the maximum difference observed in the entire table. The differences between mode 2 for every arrhythmia type is captured in the bottom matrix. The abbreviations for each arrhythmia type are given in Table 1 in the Appendix.

Mode 3

Symbols	!	(AFIB	(AFL	(B	(BII	(IVR	(NOD	(SVTA	(T	(VFL	(VT	A	E	F	L	N	R	S	V	x
!	0																			
(AFIB	0.114	0																		
(AFL	0.209	0.144	0																	
(B	0.068	0.06	0.15	0																
(BII	0.191	0.192	0.175	0.173	0															
(IVR	0.056	0.098	0.187	0.061	0.157	0														
(NOD	0.101	0.079	0.143	0.071	0.15	0.071	0													
(SVTA	0.199	0.099	0.129	0.142	0.187	0.165	0.124	0												
(T	0.168	0.062	0.125	0.106	0.204	0.141	0.108	0.064	0											
(VFL	0.245	0.169	0.157	0.197	0.154	0.204	0.171	0.092	0.146	0										
(VT	0.093	0.06	0.129	0.046	0.162	0.073	0.034	0.122	0.095	0.177	0									
A	0.122	0.062	0.106	0.063	0.146	0.1	0.079	0.098	0.083	0.143	0.059	0								
E	0.204	0.106	0.08	0.142	0.196	0.178	0.126	0.066	0.076	0.125	0.116	0.095	0							
F	0.197	0.108	0.102	0.134	0.209	0.183	0.158	0.106	0.093	0.155	0.135	0.089	0.084	0						
L	0.078	0.069	0.153	0.035	0.173	0.08	0.094	0.148	0.117	0.196	0.069	0.06	0.146	0.123	0					
N	0.173	0.08	0.088	0.111	0.167	0.147	0.109	0.06	0.065	0.113	0.096	0.055	0.05	0.062	0.109	0				
R	0.171	0.076	0.104	0.109	0.186	0.151	0.125	0.076	0.066	0.133	0.105	0.06	0.07	0.04	0.102	0.032	0			
S	0.215	0.187	0.148	0.18	0.1	0.194	0.178	0.174	0.194	0.144	0.176	0.135	0.171	0.159	0.167	0.139	0.15	0		
V	0.105	0.056	0.128	0.043	0.179	0.098	0.096	0.128	0.092	0.184	0.066	0.046	0.121	0.095	0.037	0.085	0.075	0.165	0	
x	0.087	0.032	0.148	0.037	0.176	0.072	0.06	0.116	0.087	0.176	0.04	0.056	0.122	0.124	0.051	0.093	0.093	0.179	0.051	0

Mode 4

Symbols	!	(AFIB	(AFL	(B	(BII	(IVR	(NOD	(SVTA	(T	(VFL	(VT	A	E	F	L	N	R	S	V	x
!	0																			
(AFIB	0.06	0																		
(AFL	0.293	0.241	0																	
(B	0.04	0.027	0.257	0																
(BII	0.223	0.18	0.101	0.191	0															
(IVR	0.118	0.094	0.201	0.095	0.114	0														
(NOD	0.085	0.04	0.212	0.047	0.147	0.068	0													
(SVTA	0.106	0.069	0.195	0.075	0.126	0.041	0.043	0												
(T	0.124	0.07	0.174	0.085	0.123	0.078	0.043	0.047	0											
(VFL	0.193	0.148	0.135	0.164	0.101	0.114	0.13	0.103	0.103	0										
(VT	0.053	0.033	0.243	0.019	0.175	0.08	0.033	0.061	0.074	0.154	0									
A	0.088	0.044	0.209	0.056	0.142	0.056	0.029	0.029	0.046	0.112	0.046	0								
E	0.088	0.053	0.237	0.071	0.18	0.102	0.074	0.08	0.087	0.122	0.075	0.057	0							
F	0.101	0.049	0.197	0.066	0.141	0.073	0.036	0.038	0.033	0.102	0.058	0.023	0.057	0						
L	0.055	0.036	0.261	0.046	0.198	0.106	0.07	0.087	0.098	0.153	0.057	0.062	0.038	0.069	0					
N	0.115	0.063	0.185	0.08	0.129	0.07	0.047	0.038	0.034	0.088	0.072	0.031	0.061	0.017	0.079	0				
R	0.085	0.045	0.227	0.062	0.166	0.084	0.059	0.061	0.073	0.115	0.063	0.038	0.02	0.041	0.039	0.046	0			
S	0.247	0.214	0.242	0.232	0.222	0.211	0.219	0.201	0.206	0.133	0.232	0.195	0.163	0.188	0.197	0.178	0.171	0		
V	0.068	0.027	0.236	0.042	0.175	0.088	0.047	0.061	0.071	0.131	0.045	0.037	0.033	0.041	0.031	0.052	0.023	0.192	0	
x	0.032	0.044	0.266	0.019	0.196	0.094	0.056	0.08	0.097	0.172	0.024	0.065	0.082	0.078	0.057	0.091	0.073	0.241	0.054	0

Figure 11: The differences between mode 3 for every arrhythmia type are captured in the top matrix. The differences between mode 4 for every arrhythmia type are captured in the bottom matrix. The abbreviations for each arrhythmia type are given in Table 1 in the Appendix.

Figure 12: The differences between mode 5 for every arrhythmia type are captured here. The abbreviations for each arrhythmia type are given in Table 1 in the Appendix.