

Feature Extraction of ECG signals using Wavelet transform ECE 251C Final Project report

ECE 251C Final Report Submission

Group 7

Abstract

The ABSTRACT is to be in fully justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word "Abstract" as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. Leave two blank lines after the Abstract, then begin the main text. Look at previous CVPR abstracts to get a feel for style and length.

1. Introduction

In today's age Rising Cardiovascular diseases pegs need of timely diagnosis. Apart from regular health checkups an easy & accessible heart health diagnosis solution is needed to prevent mishaps like heart failure. One such solution can be - to use the pulse data (similar to ECG data) from wearable health monitoring devices like fitbit and analyze that data to detect/diagnose an early warning sign which will enable us to predict the onset of a cardiovascular disease. There are various factors which we would have to incorporate such a system - like the power consumption, performance, precision of measurement, desired accuracy of the predictions, etc. Now, to build a full ECG abnormality detection and classification mechanism, we would have to choose and implement proper techniques to perform the following actions- data preprocessing, feature extraction, and classification. In this project, our main focus would be on the feature extraction step of this design and the usage of wavelet decomposition and analysis techniques - namely - Discrete Wavelet Transform (DWT) and Wavelet Packet Decomposition (WPD) for the feature extraction of ECG signals. To see how effective DWT and WPD feature extraction techniques are with respect to different classifiers and what impact do they have on classification accuracy, we have performed the experiment where we vary the scale(5,9,10), mother wavelet type (db4, db6, haar, symlet4), technique(DWT/ WPD) and classifier(SVM, KNN,

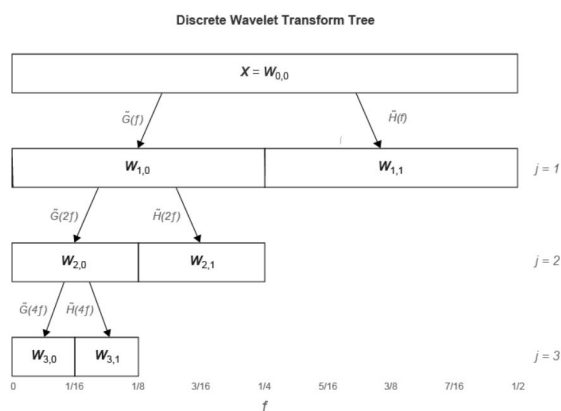
Fine decision Tree) to obtain the best working combination of these to maximize the accuracy of the system.

2. Previous Works

ECG signals have been traditionally been used to identify cardiovascular abnormalities present in the human body. However, pure ECG signals are hard to come by due to the fact that breathing pattern frequencies or electrical interferences corrupt the signal to a great extent. As a result, traditional approaches have relied on several preprocessing to actually make the signal worthy of inputting it further into data processing systems. Two types of abnormalities corrupt the ECG signals, the first one being the baseline wandering and the next one being noise.



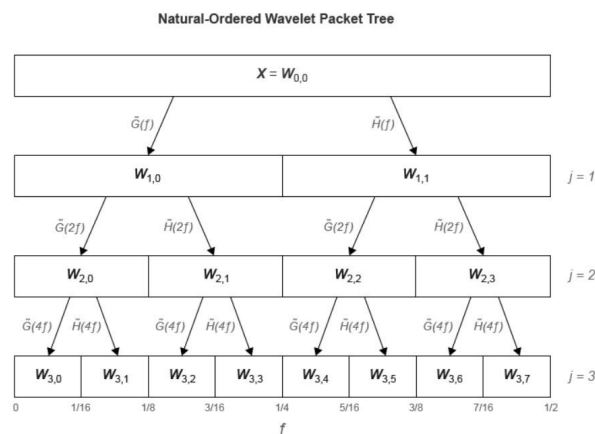
Baseline wandering is an artifact present in an ECG signal and the causes for baseline wandering are the respiration of the patient and movement while recording the ECG. These make the signal hard to interpret as there might be additional frequencies other than the desired heart rate frequency alone. The traditional method of removing the baseline wandering is passing the signal into a high pass filter. 111 paper has a section that talks about many such filtering methods for removing baseline wandering. Noise in an ECG signal creeps into the ECG recording if the recording pads are not kept properly, Electromyographic noise, and power line interferences. These noises can be removed by utilizing a low pass notch filter at 50/60Hz usually. But [Cite 1] also suggests further developments that have been used to remove the noise more efficiently. Once the preprocessing is done, different ways are used to extract features which is then fed into the classifier. A classic feature extraction method is the Fourier transform method, peak detection(P-QRS-T wave detection). Except for the Fourier transform technique, other techniques mainly rely



on time-based abnormality detections. In the Peak detection technique, the distance between the peaks acts as the main classification basis which is computed as the distance between two R points in the P-QRS-T wave [cite 2]. A simple method of classification is the Principle component analysis of the signal itself though which the classifier will handle by itself. For frequency-based features, performing the FFT followed by the classification based on FFT coefficients is traditionally simple and gives the frequency-based feature classification.

3. Proposed Approach

In our proposed approach, we use DWT and WPD to perform the feature extraction step instead of using conventional feature extraction methods for ECG signals. The concept of wavelet transform can be explained simply as decomposing or transforming a signal into multiple orthogonal basis functions. However, the difference between Fourier transform and wavelet transform or decomposition is the fact that Fourier provides resolution in frequency domain alone whereas wavelets act like a hierarchical setup of basis functions across time, where we can point out which frequency is present in what time to a considerable extent. As shown in the figure, DWT decomposes the signal into high pass and low pass sections and as the scales go beyond 1, it takes the low pass functions alone to further look deep into its components. The main difference between discrete Wavelet transform and Wavelet packet decomposition stems from what these techniques decompose and look further. Wavelet packet decomposition takes both the lowpass and highpass components and decomposes them into further lowpass and highpass, thus providing a higher overall resolution of the complete signal. Thus, Wavelet packet transform results in equal-width sub-band filtering of signals as opposed to the coarser band filtering found in DWT. [1–5] Comparison experiment variables - Mother Wavelet type (Haar, DB, Symlet), Scale of DWT/WPD, Type of Decomposition (WPD or DWT).



By using these two approaches, we aim to optimize the accuracy of the classifier and also provide better (workable) prediction results for low-computational classifiers. Apart from this, feature extraction using this method also eliminates the need for noise removal in the data pre-processing step, thus, we can say that using this technique still keeps the classification process immune to non-linearities of the signal even after removing the noise filtering process in the data pre-processing step.



The mentioned Non-linearities of ECG signals like powerline interference will not impact the classification result while using the wavelet transform for feature extraction because when we select the features obtained from wavelet decomposition, we would ignore the data present in trivial frequency bands (containing non-essential data/ noise) and we would not feed that non-essential data to the classifier, thus denoising can be done at the feature extraction step itself. Once we are done with the basic data pre-processing like baseline wandering removal, we need to compute the DWT/WPD coefficients for all scales. After calculating the DWT approximate and detailed coefficients, we use the selected scales of data that are best suited for our classification requirements. We take the coefficients of the selected scale and we divide the coefficients' data into smaller chunks and calculate the energy content of each such chunk. Once we have the energy values of these chunks, we calculate the distribution of energy in these chunks and give that as an input parameter to the classifier. The classifier then takes these parameters as input from every such scale which was selected by us and runs the respective classification algorithm. We then compare the obtained results from the ones obtained from the traditional classification methods. In WPD, for feature extraction purposes, 333 had a good idea about

using all the coefficients of the last scale of decomposition. Using this idea as the base, we checked the energy percentage distribution across coefficients and found that most of the energy was concentrated around 1-16 indexes. We have used these distribution parameters as our feature extraction pivots and have classified the healthy and unhealthy data sets. We will perform a comparative study for determining which type of mother wavelet is best suited for this application and up to what scale (levels) of cascading DWT/ WPD is useful for this process. furthermore, we would also check if the features extracted using wavelet packet decomposition (WPD) give better results than DWT.

4. Simulation Results, Comparison & Discussion

4.1. Simulation Environment

The simulation environment consisted of us extracting the data from [repo link] which consist of patients with various heart arrhythmia. Each data was of length 30 mins. So in order to reduce the computationally convenient and add more data, we chose to take individual sections of the ECG that had problems as our unhealthy signal and portions of ECG signal that did not have any arrhythmia as healthy ones. Each of the unhealthy and healthy signal length lasted for a minute. We then used the input signal after doing pre-processing to get the features that we have mentioned above and fed it into the classifiers to see the performance.

4.2. Simulation Results

We ran the simulation for 5 types of wavelets and depending on the type of method used DWT/WPD, we chose scales. Scales of 9 and 10 were used for DWT and scales of 5 and 9 were used for WPD. The tables X and Table Y shows us the accuracy results of the classification for a particular type of classifier.

KNN Classifier	Scale : 9		Scale : 10	
Method	Mother Wavelet	Accuracy	Mother Wavelet	Accuracy
DWT	Db2	80	Db2	86
	Db4	72	Db4	79
	Db6	75	Db6	84
	Sym4	80	Sym4	81
	Haar	72	Haar	70

As we can see from the results, in general of performance of WPD is better than DWT which is expected because of the fact that WPD will have more information when compared to the DWT of the same scale.

KNN Classifier	Scale : 5		Scale : 9	
Method	Mother Wavelet	Accuracy	Mother Wavelet	Accuracy
WPD	Db2	80	Db2	86
	Db4	78	Db4	84
	Db6	84	Db6	88
	Sym4	80	Sym4	82
	Haar	78	Haar	78

SVM Classifier	scale 9	scale10		scale 5	scale9
DWT	Accuracy	Accuracy	WPD	Accuracy	Accuracy
Mother Wavelet Db6	81.4	85.3	Mother Wavelet Db6	83	87

Decision Tree Classifier	scale 9	scale10		scale 5	scale9
DWT	Accuracy	Accuracy	WPD	Accuracy	Accuracy
Mother Wavelet Db6	72	84	Mother Wavelet Db6	82	83

5. Conclusion

In conclusion, we can deduce that the proposed method of feature extraction using DWT and DWD is much more computationally efficient (large scale data) than conventional methods (peak detection, QRS Detection), and by using this method we are able to reduce the number of features/dimensions of the data being fed to the classifier, thus it drastically reduces the latency of the classifier. Using wavelet decomposition also doesn't negatively affect the amount of useful information sent to the classifier that is used for classification. We were also able to improve the accuracy of the classification results to a good extent. We con The filtering step in the pre-processing stage is being done at the feature extraction step itself instead of using separate stages. This resulted in the system being more time-efficient and the complexity of the overall system was hence reduced. Further, we can say that this method can be easily applied in multi-class classification for detecting different types of abnormalities like arrhythmia, atrial flutter, fibrillation, etc to make the diagnosis more precise.

References

- [1] A. Senthilkumar M. Sasireka. Dwt based feature extraction from ecg signal and detection of cardiac diseases. *IJAER*, 2015. 2
- [2] Taiyong Li Min Zhou. Ecg classification using wavelet packet entropy and random forests. *Entropy*, 18(285), 2006. 2
- [3] Risanuri Hidayat Siti Agrippina Alodia Yusuf. Feature extraction of ecg signals using discrete wavelet transform and mfcc. *IEEE ICSITECH*, 12(1):167–170, 2019. 2
- [4] Dr. Devendra Prasad V.K.Srivastav. Dwt - based feature extraction from ecg signal. *American Journal of Engineering Research (AJER)*, 2:44–50, 2013. 2

- [5] Shehzad Khalid Waqas Ahmed. Ecg signal processing for recognition of cardiovascular diseases: A survey. pages 677–682, 2016. 2

Miscellaneous Links

- [Open Source Database](#)
- [Repository](#)