Amplitude and Phase Classification of ECG data

Chibueze Emmanuel Ogbonnaya
With: S. P. Preston, A. T. A. Wood and K. Bharath

University of Nottingham

pmxceog@nottingham.ac.uk

September 2, 2018

Overview

Electrocardiogram

- An Electrocardiogram (ECG) is used to record the electrical activity of the heart to identify and locate pathology.
- The ECG is essential for the diagnosis and management of abnormal cardiac rhythms.

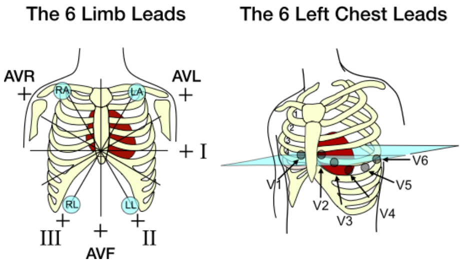
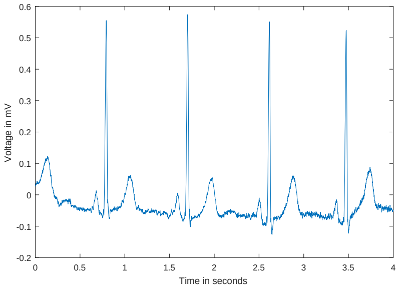
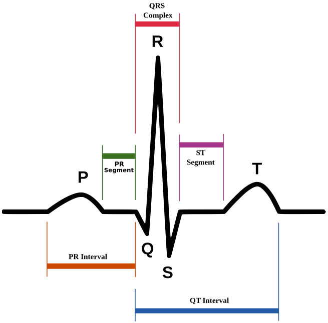


Figure: ECG 12-lead placement

ECG cont'd.



(a) ECG Signal for an individual from lead I

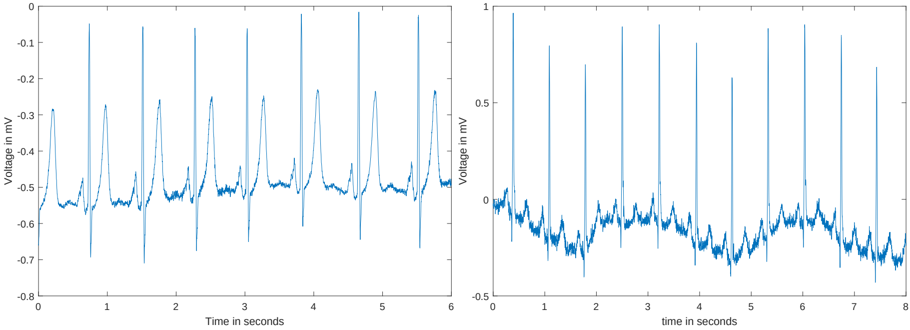


(b) Features of a Typical ECG Signal

Figure: Sample ECG Signal and Features

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

It is difficult to detect ECG features in a noisy ECG signal.

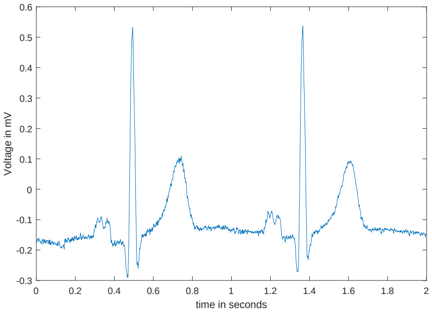


(a) Nice ECG signal (b) Noisy ECG signal with baseline drift

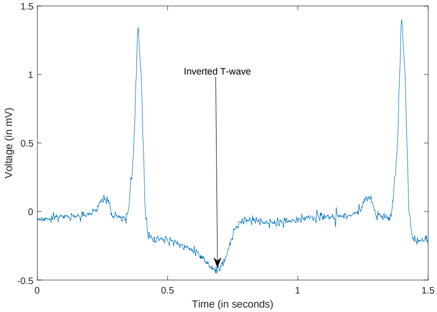
Figure: ECG Signals of two different Healthy Controls showing effect of Noise

Healthy Controls and Heart Conditions

- Myocardial Infarction: ST elevation
- Cardiomyopathy: inverted T wave and prolonged QT interval.



(a) Normal ECG for healthy control



(b) ECG with inverted T wave

Figure: ECG changes caused by heart conditions

We propose using RR intervals for functional representation of an ECG signal.

- An RR interval corresponds to a heartbeat.
- The RR interval contains important features of interest: ST segment and T wave.

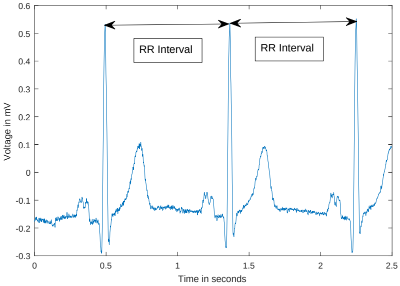
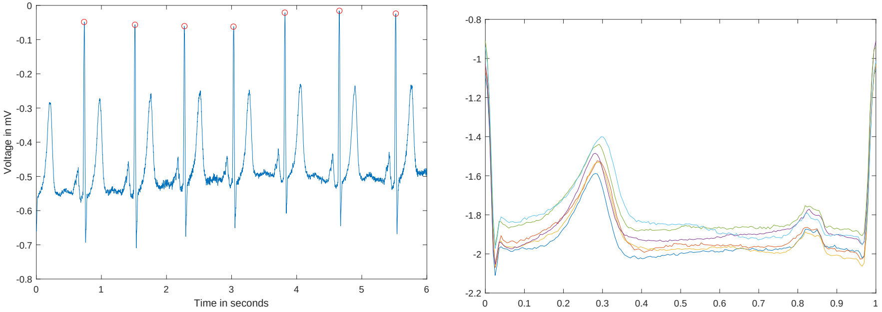


Figure: An ECG signal showing RR intervals



(a) ECG with R peaks shown in Red

(b) Chopped-up RR functions

Figure: ECG signal and chopped-up functions

Data Source

Data is taken from the PTB ECG database and contains 10 seconds recordings. We consider the conventional 12-leads.

For our data, we have ECG for

- Healthy controls
- Myocardial infarction
- Cardiomyopathy.

Target

Registration and Fitting

- Remove noise in ECG signals through amplitude registration.
- Estimate amplitude and phase components of registered ECGs using parametric models.

Classification

Classification of ECGs using estimated amplitude and phase components.

Model

$$y_i(t) = b_i x_i(t) + \sum_{j=1}^{q} a_{ji} u_j(t)$$

where $t \in [0, 1]$. We define the following:

- $x_i(t)$ are the observed RR functions.
- $y_i(t)$ are the registered RR functions
- $u_j(t)$ form an orthonomal basis function for noise.

Registration implies estimating a_{ji} and b_i with template $f(t)$

$$\underset{a_{ji}, b_i \in \mathbb{R}}{\text{minimise}} \quad \sum_{i=1}^{n} \|y_i - f\|^2.$$

Example Solution for $f(t) = 0$

- Use the zero function as template

Solutions:

$$\hat{a}_{ji} = -b_i \langle x_i, u_j \rangle,$$

$$y_i(t) = b_i \left(x_i(t) - \sum_{j=1}^{q} \langle x_i, u_j \rangle u_j(t) \right).$$

- Estimate b_i: Constraint $\sum_{i=1}^{n} \log b_i = 0$.
- $\hat{b}_i = c_i^{-1/2} (\prod_{i=1}^{n} c_i^{-1/2n})$.
- $c_i = \|x_i\|^2 - \sum_{j=1}^{q} \langle x_i, u_j \rangle^2$.

Amplitude Registration: Example

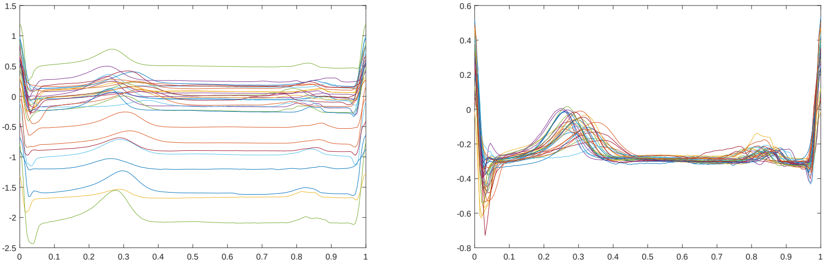


Figure: Left: Observed RR functions. Right: Amplitude registration.

- We use sum of Gaussian functions and B-splines to fit the RR functions
- The Gaussian models have been used previously to generate synthetic ECGs in Clifford (2006).

For the Gaussian mixture model with phase $t \in [0,1]$, we have

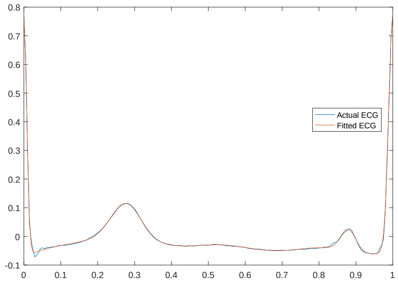
$$z(t, \boldsymbol{\alpha}_i, \boldsymbol{\theta}_i, \boldsymbol{\beta}_i) = \sum_{j=1}^{k} \alpha_{ij} \exp\left[-\frac{(t-\theta_{ij})^2}{2\beta_{ij}^2}\right] \tag{1}$$

where $\theta_{i1} = 0 \leq \theta_{i2} \leq \ldots \leq \theta_{ik} = 1$.

To fit this model to actual ECG signal $y_i(t)$, we will need to solve the non-linear optimisation problem

$$\min_{\boldsymbol{\alpha}_i,\boldsymbol{\theta}_i,\boldsymbol{\beta}_i} \quad \int_0^1 (y_i(t) - z(t,\boldsymbol{\alpha}_i,\boldsymbol{\theta}_i,\boldsymbol{\beta}_i))^2 dt$$

$$\text{subject to} \quad \theta_{i2} - \theta_{i3} < 0, \theta_{i3} - \theta_{i4} < 0, \ldots, \theta_{i(k-2)} - \theta_{i(k-1)} < 0.$$



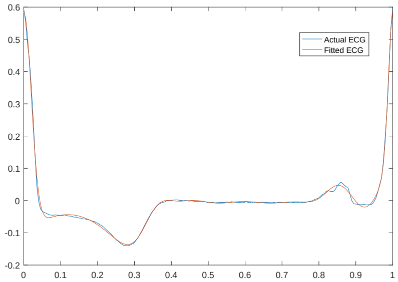


Figure: Fitting gaussian parametric models to actual ECGs ($k = 11$). Left: healthy control ECG. Right: Cardiomyopathy ECG.

1. Determine Template $\hat{\mu}(t)$.
2. Using template, fix $\boldsymbol{\beta}_i = \boldsymbol{\beta}$, $\boldsymbol{\theta}_i = \boldsymbol{\theta}$, $i = 1, \ldots, n$

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\theta}} \int_0^1 (\hat{\mu}(t) - z(t, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}))^2.$$

3. New Model: $z(t, \boldsymbol{\alpha}_i) = \sum_{j=1}^{k} \alpha_{ij} \exp\left[-\frac{(t-\theta_j)^2}{2\beta_j^2}\right]$.
4. For each i, estimate $\boldsymbol{\alpha}_i$.

1. Determine Template $\hat{\mu}(t)$.
2. Using template, fix $\boldsymbol{\beta}_i = \boldsymbol{\beta}$, $\boldsymbol{\alpha}_i = \boldsymbol{\alpha}$, $i = 1, \ldots, n$

$$\min_{\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\theta}} \int_0^1 (\hat{\mu}(t) - z(t, \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\theta}))^2 dt.$$

3. New Model: $z(t, \boldsymbol{\theta}_i) = \sum_{j=1}^{k} \alpha_j \exp\left[-\frac{(t-\theta_{ij})^2}{2\beta_j^2}\right]$
4. For each i, estimate $\theta_{i2}, \ldots, \theta_{i(k-1)}$.

Amplitude Two Sample T-test

1. We conduct a two-sample Hotelling's t-test.
2. Healthy vs Cardiomyopathy
3. NULL HYPOTHESIS: No difference in amplitude

Result

- $F = 19.3507$, p-value $= 1.4433 \times 10^{-15}$.
- Strong evidence of difference in amplitude.

Classification Results: Cardiomyopathy

Classification is done using the estimated components, for Lead I.

Accuracy		
Method	LDA	SVM
Gaussian	0.9714	0.9429
BSpline	0.9714	0.8714

Table: Amplitude Classification of Cardiomyopathy

Accuracy		
Method	LDA	SVM
Gaussian	0.9286	0.9000
BSpline	0.9571	0.9571

Table: Phase Classification of Cardiomyopathy

Classification Results: Myocardial Infarction

Classification results from Lead I.

Accuracy		
Method	LDA	SVM
Gaussian	0.8477	0.8376
BSpline	0.8325	0.8426

Table: Amplitude Classification of MI

Combining multiple leads by concatenation, this improves to

Accuracy		
Method	LDA	SVM
Gaussian	0.8782	0.9086
BSpline	0.8731	0.8832

Table: Amplitude Classification of MI (Multiple Leads)

Summary/Conclusion

- We have proposed amplitude registration models for ECG signals.
- Parametric models are a good alternative to dimension reduction techniques like FPCA.
- Variable selection possible using estimated amplitude components.
- Automation greatly improves ECG diagnosis when compared to clinicians.
- Applicable to analysis of gait data for diagnosis of Parkinson's.

Reference	Method	Result
McCabe et al. (2013)[2]	Physicians	Sensitivity: 65%, Specificity: 79%
Sun et al. (2012)[3]	ST segments using 5-order polynomial	Sensitivity: 92.3%, Specificity: 88.1%
Kurtek et al. (2013)[1]	NN (SRVF)	Accuracy: 90%
Previous work	Functional PCA	Accuracy: 92.86%
Proposed	Gaussian Model	Accuracy: 90.86%

Table: Comparison of methods for detection of myocardial infarction

Comparison: Phase Classification

Reference	**Method**	**Result**
Tucker et al. (2013)	Horizontal FPCA (SRVF)	0.9429
Tang and Müller (2008)	Pairwise Synchronisation (PACE)	0.7857
Proposed	B-Spline Model	0.9571

Table: Comparison of methods for detection of Cardiomyopathy

Thank You

References

Kurtek, Sebastian and Wu, Wei and Christensen, Gary E and Srivastava, Anuj (2013)
Segmentation, alignment and statistical analysis of biosignals with application to disease classification
Journal of Applied Statistics .

McCabe, James M and Armstrong, Ehrin J and Ku, Ivy and Kulkarni, Ameya and Hoffmayer, Kurt S and Bhave, Prashant D and Waldo, Stephen W and Hsue, Priscilla and Stein, John C and Marcus, Gregory M and others (2013)
Physician accuracy in interpreting potential ST-segment elevation myocardial infarction electrocardiograms
Journal of the American Heart Association .

Li Sun and Yanping Lu and Kaitao Yang and Shaozi Li (2012)
ECG Analysis Using Multiple Instance Learning for Myocardial Infarction Detection
IEEE Transactions on Biomedical Engineering .