Acquisition and Analysis of Biosignals DTEK0042

Biosignal analysis I

Iman Azimi, Ph.D. (Tech.)
Email: iman.azimi@utu.fi

Turun yliopisto
University of Turku

20.9.2022

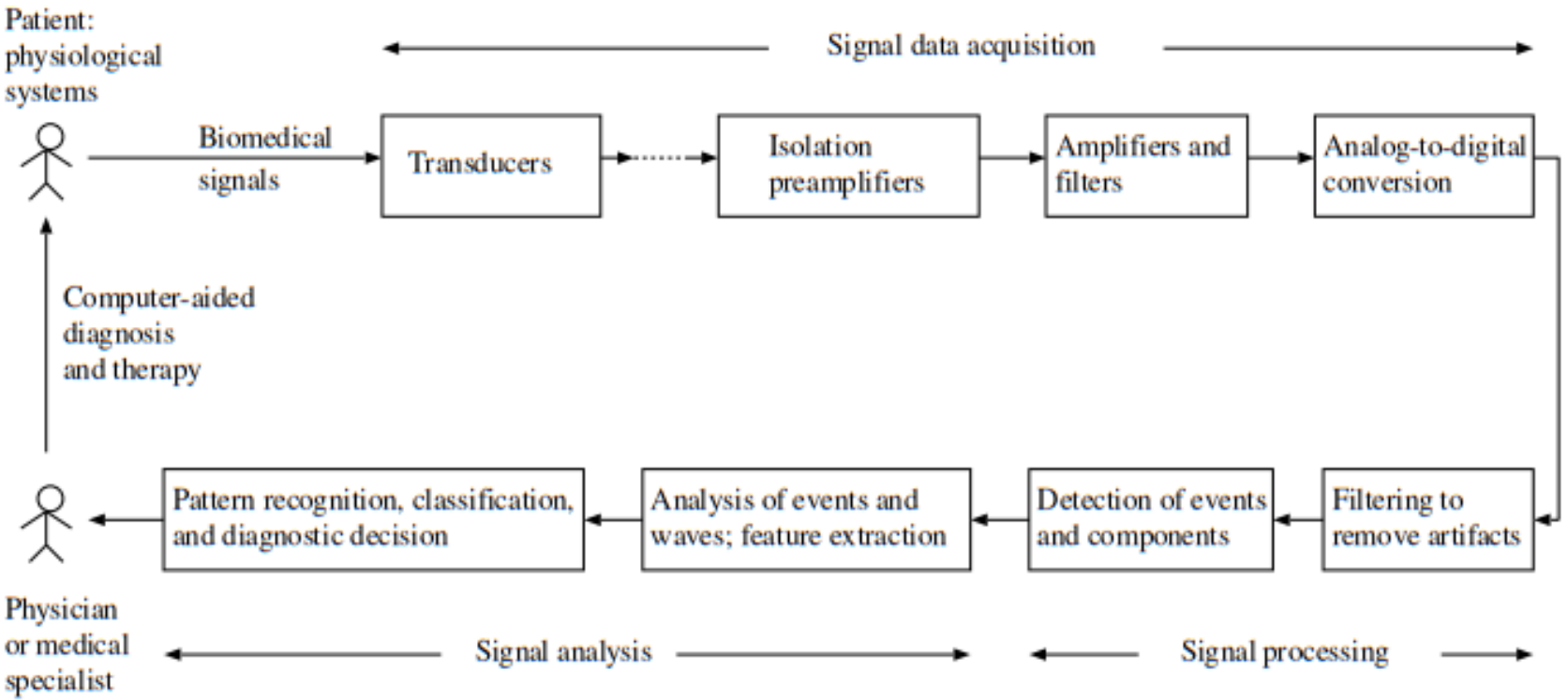
Introduction

So far, we learned:

- ❑ About the origins and acquisitions of biosignals
- ❑ Artifact removal from biosignals

In this session, we will learn:

- ❑ Objectives of biomedical signal analysis
- ❑ Correlation techniques (e.g., EEG)
- ❑ Waveform analysis and feature extraction in biosignals (e.g., ECG)



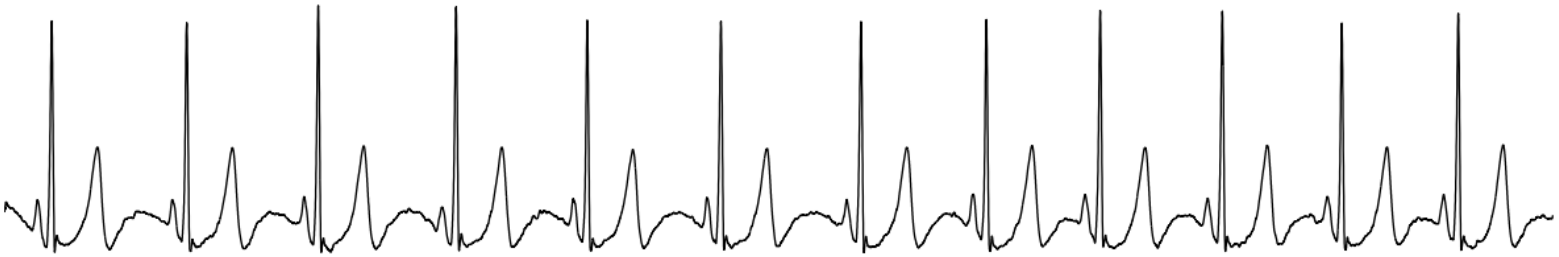
Rangayyan, R. M. Biomedical signal analysis. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Objectives of Biomedical Signal Analysis

- **Information gathering**: measurement of phenomena to interpret a system.
- **Diagnosis**: detection of malfunction, pathology, or abnormality.
- **Monitoring**: obtaining continuous or periodic information about a system.
- **Therapy and control**: modification of the behavior of a system based upon the outcome of the activities to ensure a specific result.
- **Evaluation**: objective analysis to determine the ability to meet functional requirements, obtain proof of performance, perform quality control, or quantify the effect of treatment

Biosignal analysis

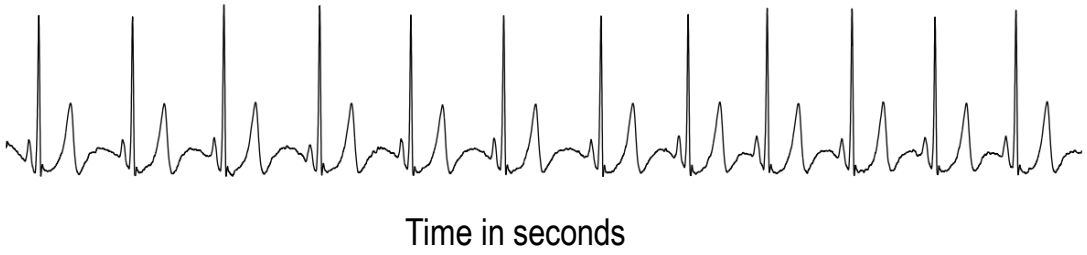
- Biomedical signals carry signatures of physiological events



- We are going to use **different time-domain and frequency-domain techniques** to analyze these events

Time-domain and Frequency-domain

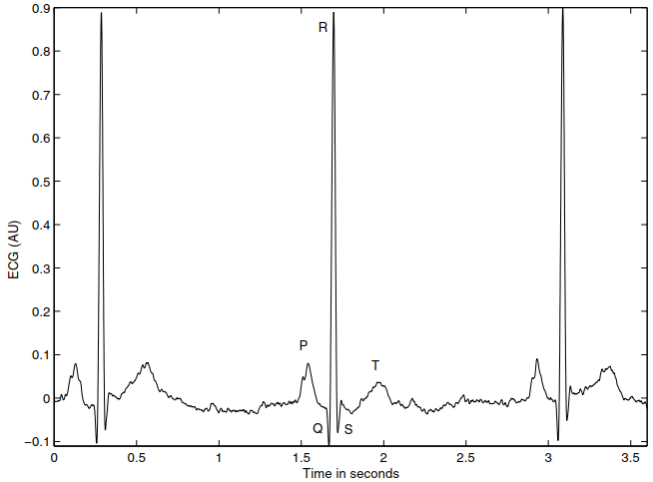
- **Fourier Transform**: decomposes a function of time (a *signal*) into its frequencies
- The events are investigated in time domain and frequency domain



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Events in biosignals

- The part of a signal related to a specific event of interest is often referred to as an epoch
- Once an event has been identified, the corresponding waveform may be segmented and analyzed in terms of:
  - Amplitude
  - Wave shape (morphology)
  - Time duration
  - Intervals between events
  - Energy distribution
  - Frequency content
  - Etc.



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Examples of events

- ECG:
  - P wave, QRS complex, and T wave
- PCG:
  - S1 and S2 sounds
- EEG:
  - Spike, Sharp wave, and α rhythm

Correlation in biosignals

Correlation in biosignals

- We are looking for the correlations and similarities in the signals:
  - If two signals are similar
  - If the signal is similar to a template
  - If there is a pattern (rhythms) in a signal

Convolution

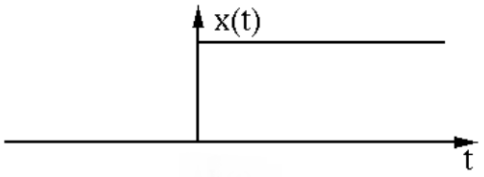
- Convolution is a mathematical operation on two functions
- Convolution is an integral that expresses the amount of overlap of one function as it is shifted over another function.

$$(x * y)(t) = \int_{\tau=-\infty}^{\infty} x(\tau)y(t-\tau)d\tau$$

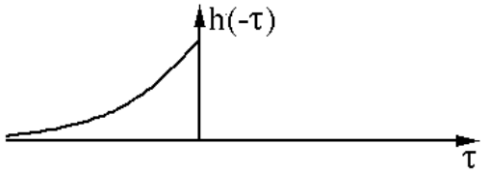
- Lets see it in some examples:

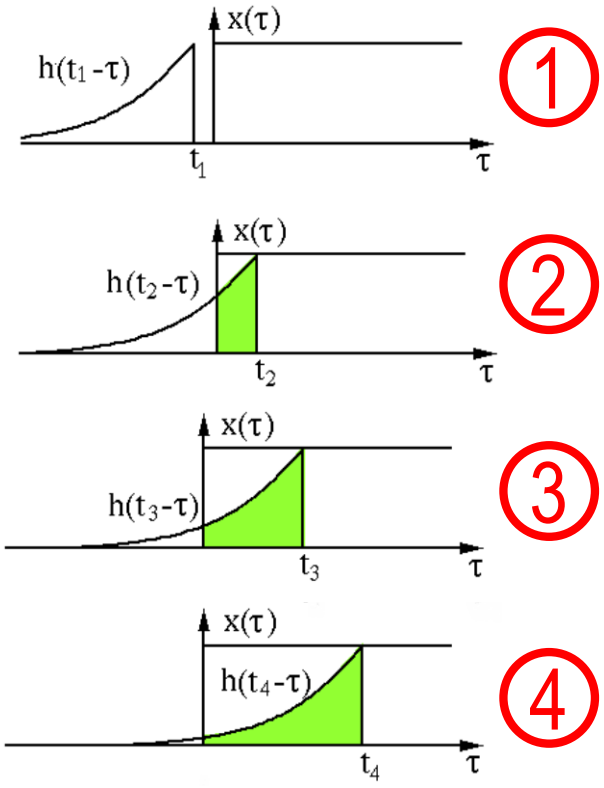
Convolution – Example 1

Function 1

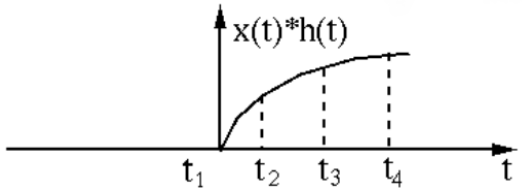


Function 2





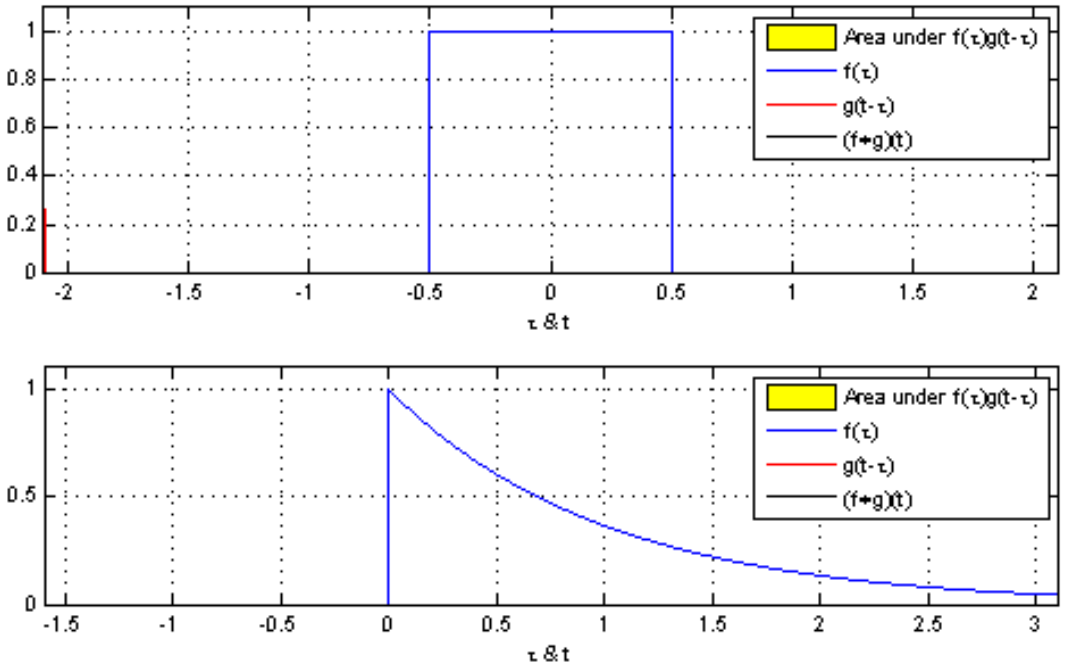
Output



http://fourier.eng.hmc.edu/e161/lectures/convolution/index.html

Convolution – Example 2

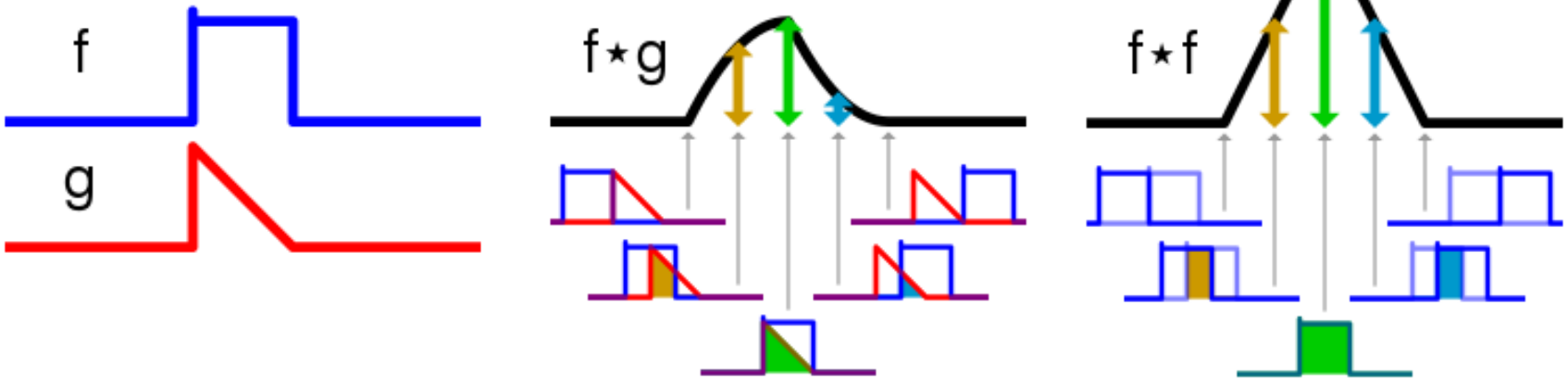
- Two examples



https://en.wikipedia.org/wiki/File:Convolution_of_box_signal_with_itself2.gif
https://en.wikipedia.org/wiki/File:Convolution_of_spiky_function_with_box2.gif

Correlations

- Peaks in the output if there are similarities



https://en.wikipedia.org/wiki/File:Comparison_convolution_correlation.svg

Cross-correlation

- Cross-correlation is the convolution of two signals

$$\theta_{xy}(\tau) = \int_{\tau=-\infty}^{\infty} x(\tau)y(t+\tau)d\tau$$

$$\theta_{xy}(k) = \sum_{n} x(n)y(n+k)$$

- Cross-correlation is a measure of similarity of two signals
- Cross-correlation displays peaks at the period of any periodic pattern present in both of the signals being analyzed
- **Cross spectral density (CSD)** is the Fourier transform of the cross-correlation => shows the strength of the energy shared by a given frequency for the two signals

Use of cross-correlation

- Cross-correlation is used to detect rhythms present in common between two signals
  - E.g: two EEG channels
- The procedure is known as **template matching**, when one of the functions being used to compute the cross-correlation is a template of an event (ECG cycle or EEG spike-and-wave complex)
  - E.g. Detection of EEG rhythms related to seizure
  - E.g. Detection of QRS complex abnormality in ECG signal
- Cross-correlation is also used to align two signals from two different sources
  - The delay between the two signals can be detected.
  - E.g.: Heart rate signals (from a person) are extracted from ECG and PPG.
  - E.g.: Respiration rate signals extracted from different sources
  - E.g.: Synchronization in the synchronized averaging filter

Auto-correlation

- Auto-correlation is the convolution of a signal with itself

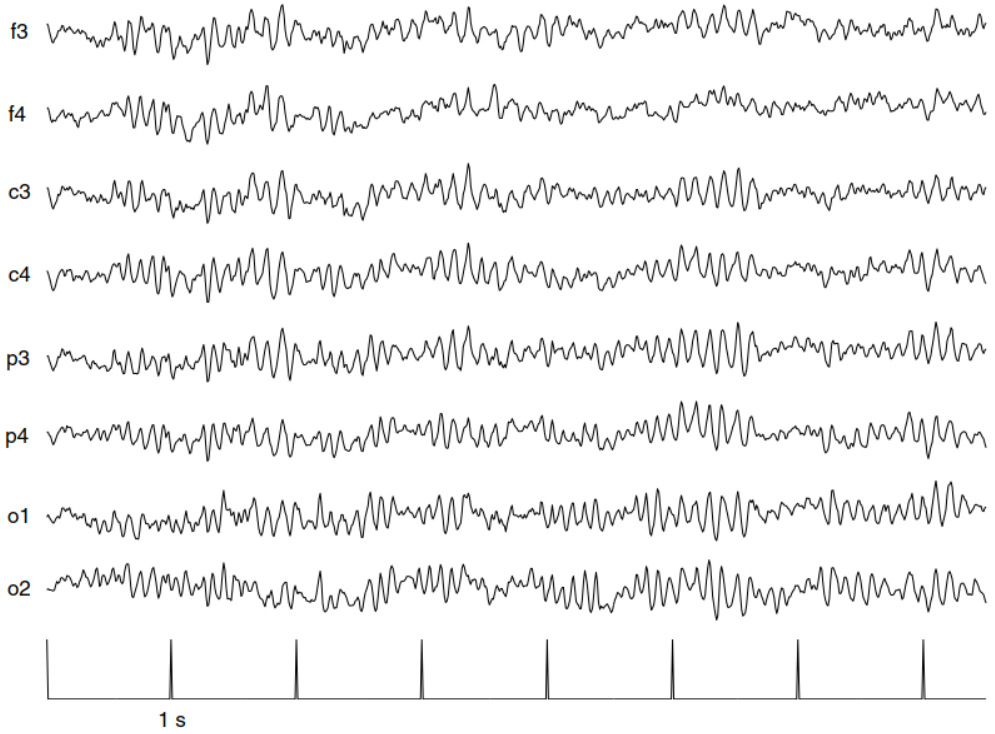
$$\theta_{xx}(\tau) = \int_{\tau=-\infty}^{\infty} x(\tau)x(t+\tau)d\tau$$

$$\theta_{xx}(k) = \sum_{n} x(n)x(n+k)$$

- Auto-correlation displays peaks at intervals corresponding to the of any periodic or repetitive pattern present in the signal.
- This property facilitates the detection of rhythms in signals
  - E.g., α rhythm in EEG signal
- Auto-correlation of most signals decays, except the periodic signals.
- **Power spectral density (PSD)** is the Fourier transform of the auto-correlation =>
  - shows the strength of the energy of the signal as a function of frequency (at which frequencies variations are strong and at which frequencies variations are weak)

EEG: α rhythm detection

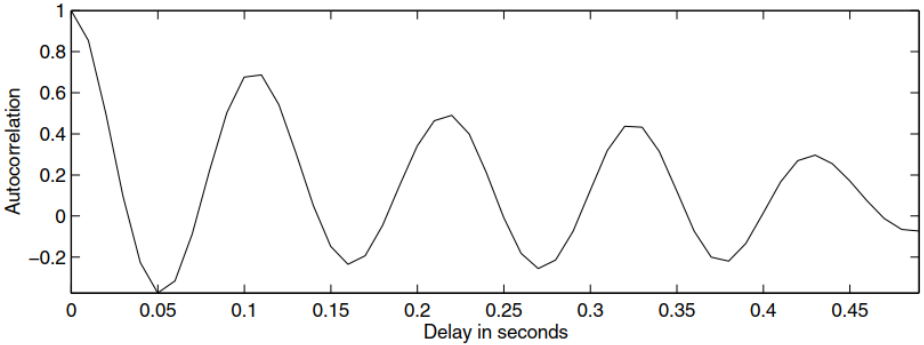
- Eight channels of the EEG
- We are going to investigate the α rhythm in the signals

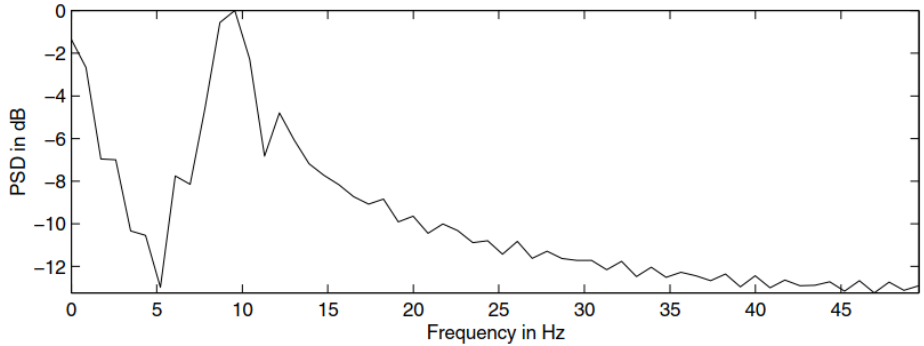


Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Example 1

- Auto-correlation of p4 over the time interval 4.67 − 5.81 s
- It shows the α rhythm with 9Hz
- Peak in PSD
- A simple peak detection can be used to detect peaks in Autocorrelation signal

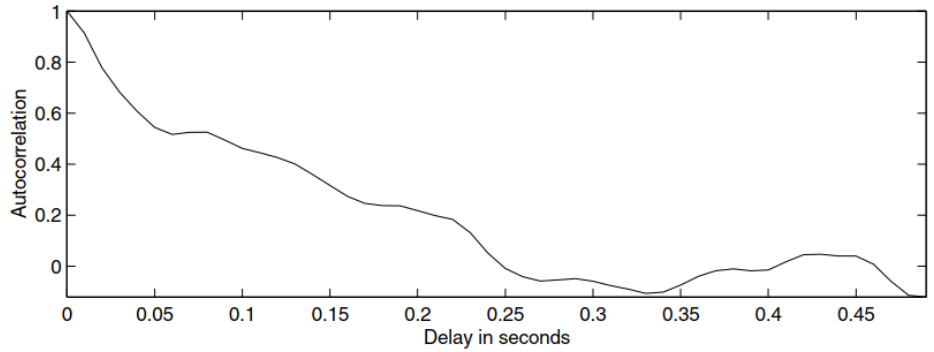


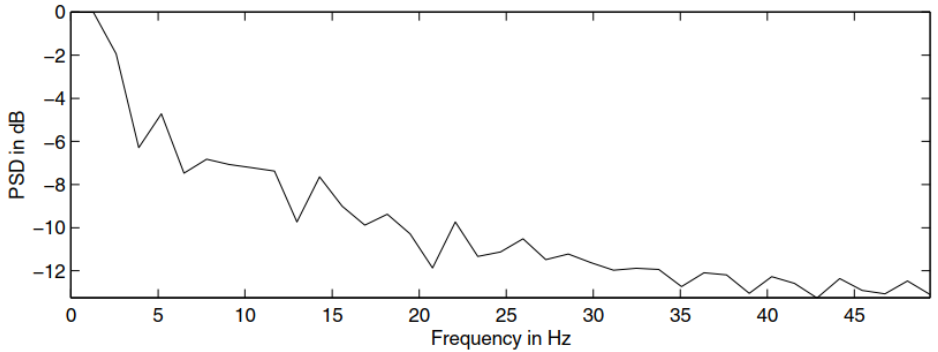


Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Example 2

- Auto-correlation of f3 over the time interval 4.2 – 4.96s
- Absence of α rhythm
- No peak in PSD

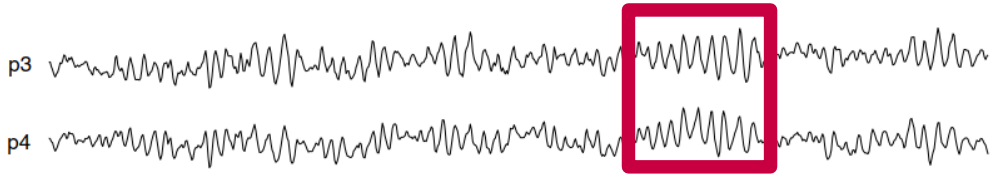


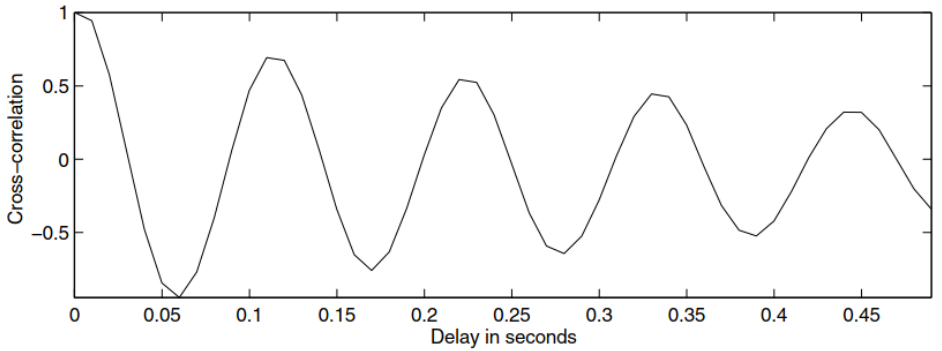


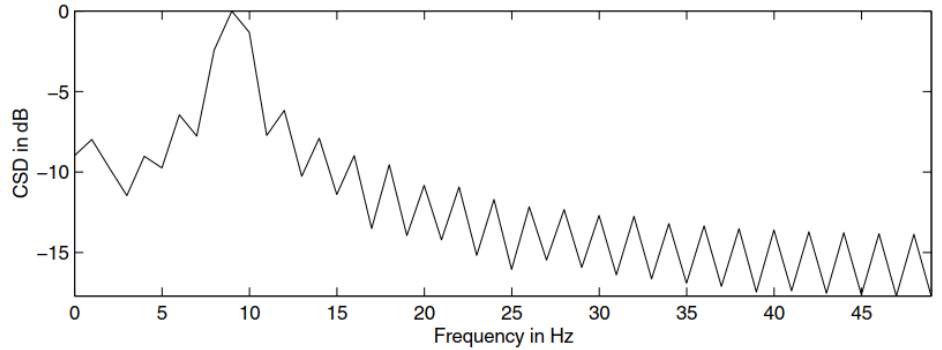
Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Example 3

- Cross-correlation of p3 and p4 over the time interval 4.72 − 5.71 s
- Both have α rhythm
- Peak in CSD



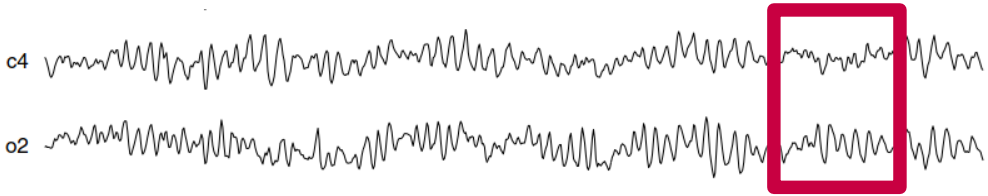


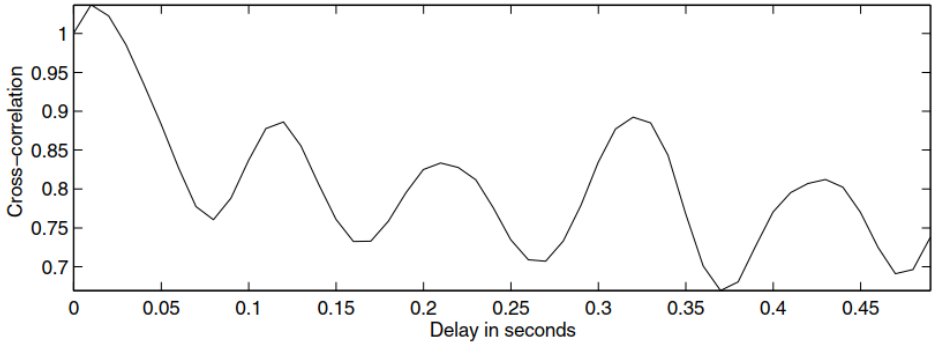


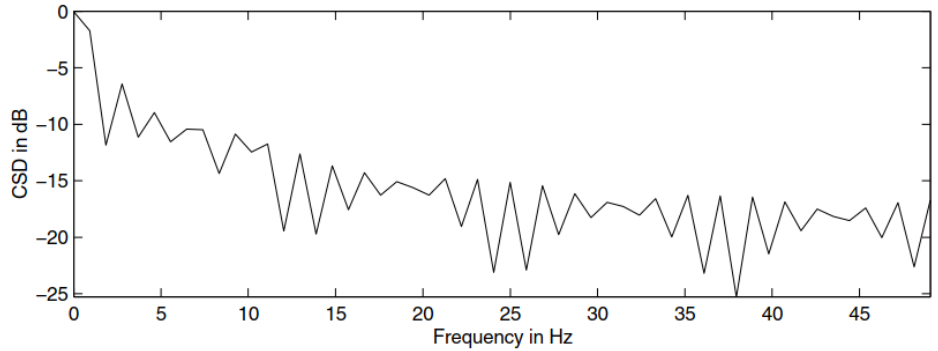
Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Example 4

- Cross-correlation of o2 and c4 over the time interval 5.71 − 6.78 s
- o2 has α rhythm but not c4
- No peak in CSD



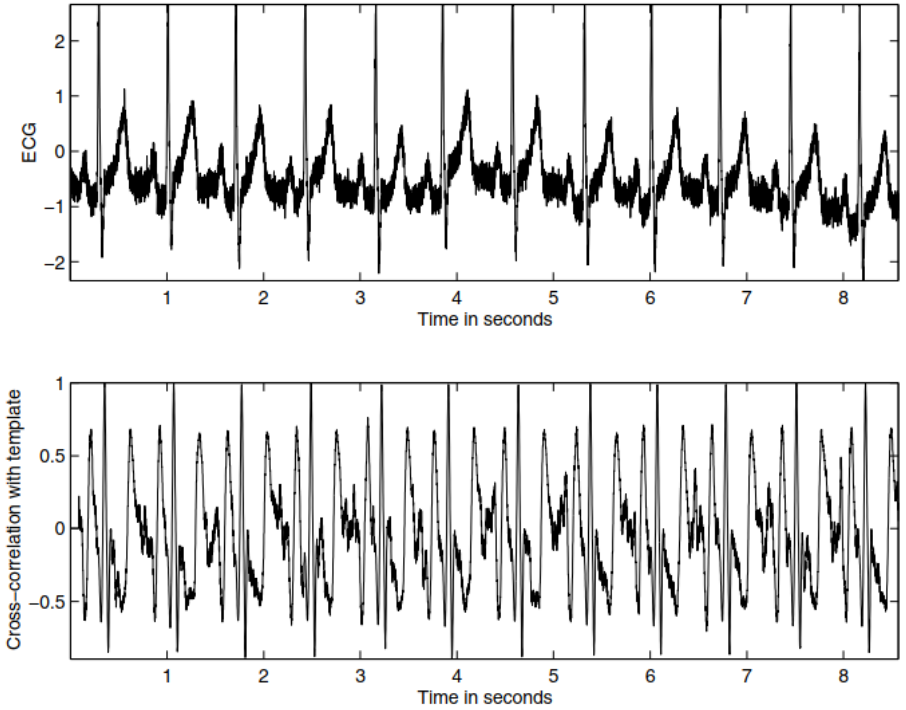




Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Example: Template matching

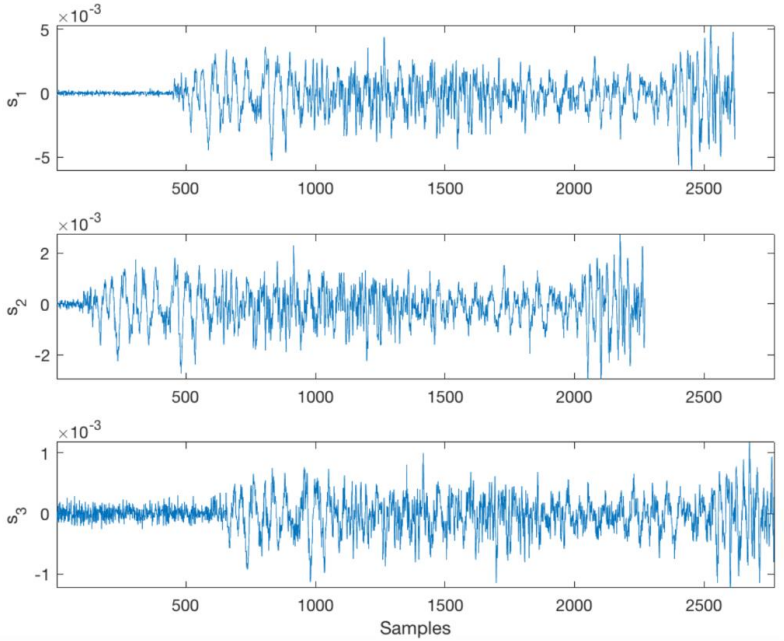
- Cross-correlation of a template and a noisy ECG
- The template is the first heart cycle
- The QRS location can be extracted via an appropriate threshold (e.g., 0.9 in this example)
- The output was normalized to [-1,1]



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

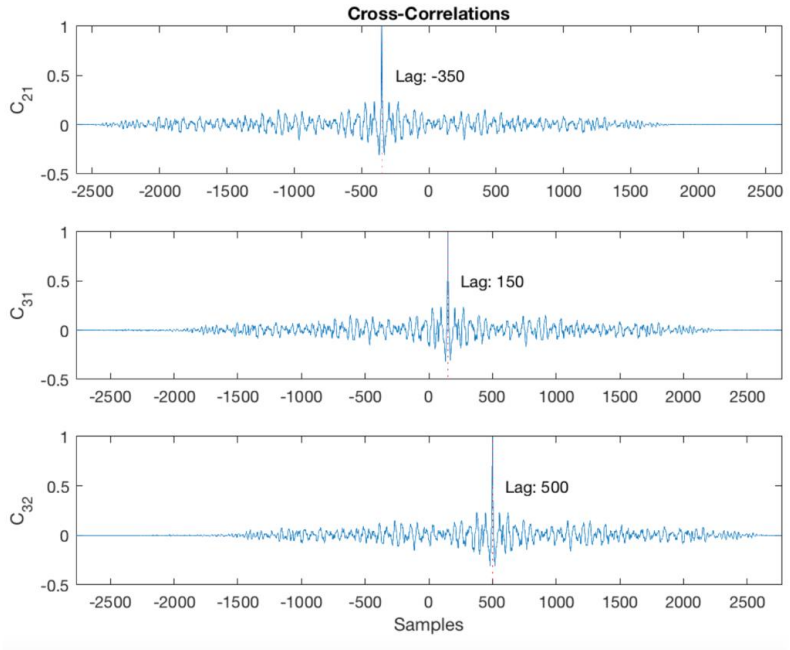
Example: alignment in signals (1)

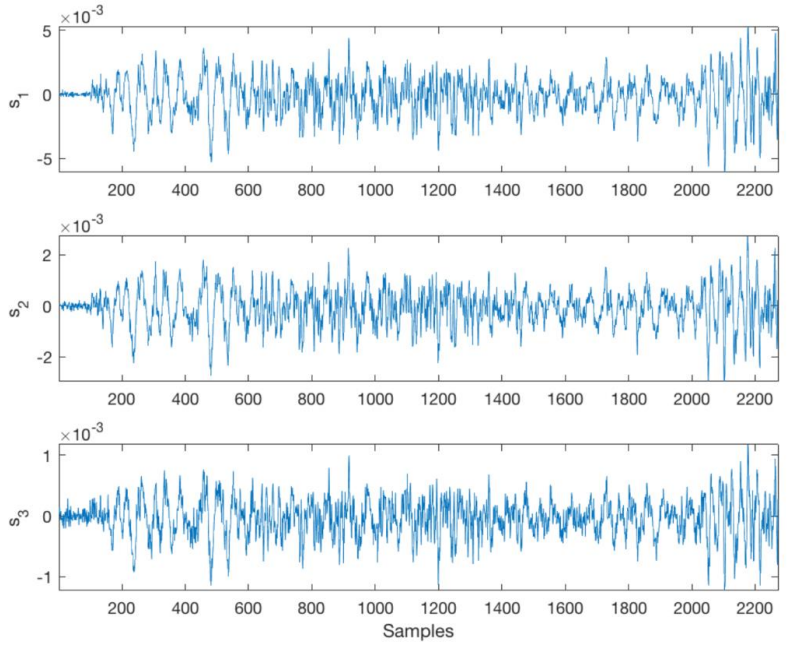
- There are 3 signals from 3 sources, which there are lags in 2 signals
- Alignment is required, before any other analysis



https://www.mathworks.com/help/signal/ug/align-signals-using-cross-correlation.html

Example: alignment in signals (2)





https://www.mathworks.com/help/signal/ug/align-signals-using-cross-correlation.html

Waveform analysis and feature extraction of ECG

Segmentation

- The signals are nonstationary
- They are repetitive including patterns
- The patterns include events
  - E.g., QRS complex
- To analyze the signal, we need to extract the patterns and divide it into the events (segments)
- We are focusing on ECG signals, however, the concept can be used in other signals as well

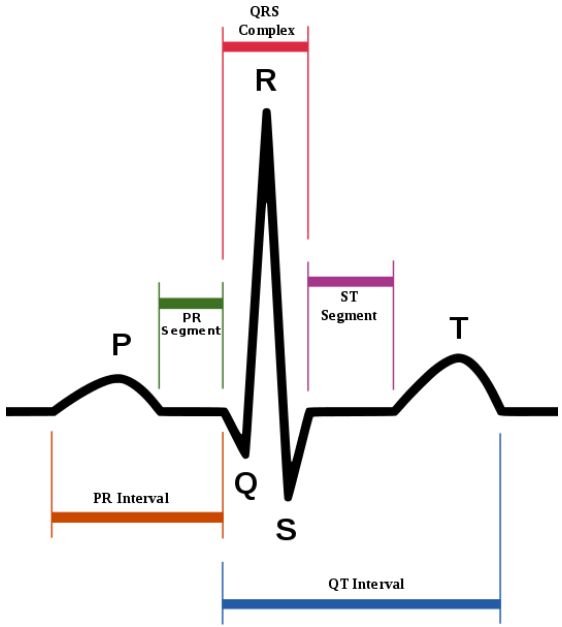
Luz, Eduardo José da S., et al. "ECG-based heartbeat classification for arrhythmia detection: A survey." Computer methods and programs in biomedicine 127 (2016): 144-164.

QRS detection in ECG

- To segment the heart cycles in ECG, QRS complex detection is required.
- QRS complex has the largest slope in a cardiac cycle.
- Two techniques to detect the QRS complex:
  - Derivative-based method
  - The Pan-Tompkins algorithm

Derivative-based method

- QRS complex has the largest slope in a heart cycle
- Derivative operator ($\frac{d}{dt}$) would be the most logical starting point
- However, the results of the derivative-based operators are noisy
- So, smoothing techniques are required to smooth the results



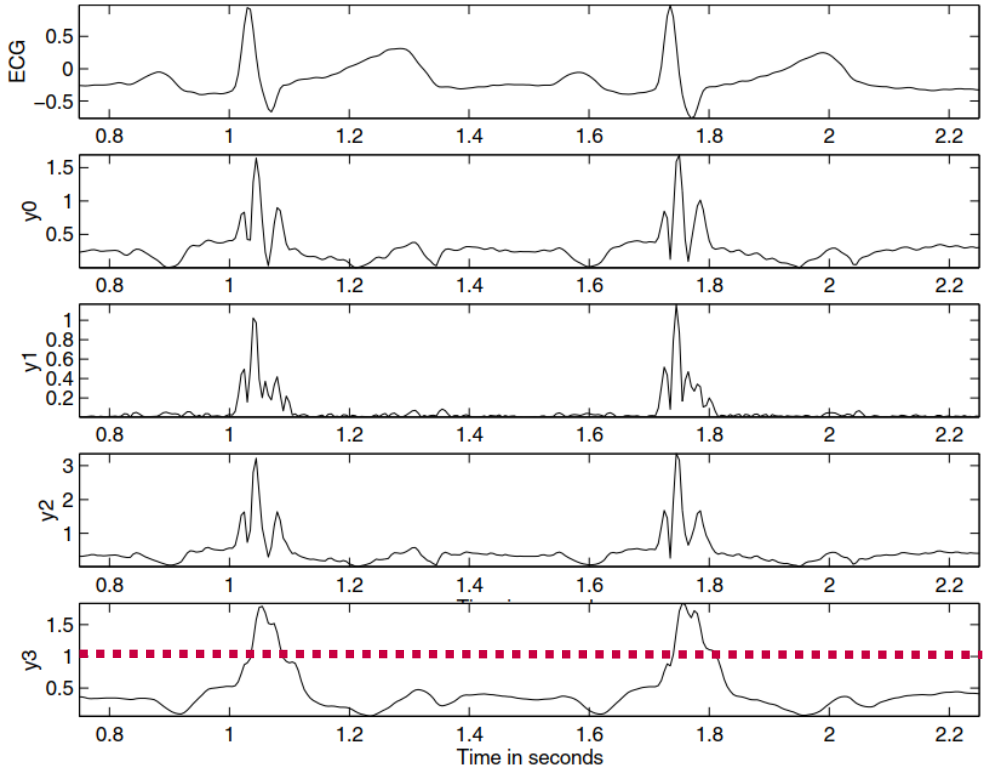
Balda et al. algorithm[1]

- The derivative-based algorithm progresses as follows:
  1. The smoothed three-point first derivative is approximated as:
  $$y_0(n) = |x(n) - x(n-2)|$$
  2. The second derivative is approximated as:
  $$y_1(n) = |x(n) - 2x(n-2) + x(n-4)|$$
  3. The first derivative and second derivative are combined:
  $$y_2(n) = 1.3y_0(n) + 1.1y_1(n)$$
  4. $y_2(n)$ is smoothed: $y_3(n)$
  5. $y_3(n)$ is scanned to check if the threshold of 1 is crossed (a peak detection method)
  6. The subsequent eight samples are also tested against the threshold. If at least six of the eight points pass the threshold test, the segment of eight samples is taken to be a part of a **QRS complex**

[1] Balda, R. A., et al. "The HP ECG analysis program." Trends in computer-processed electrocardiograms 197 (1977): 205.

Example

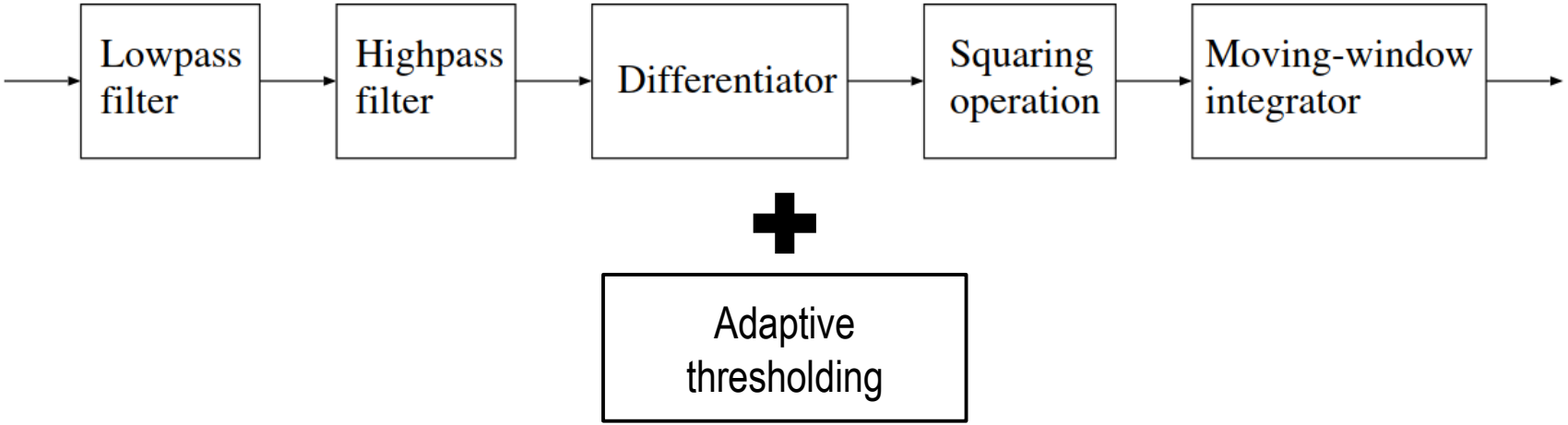
- The sampling rate is 200Hz
- The ECG signal was filtered by
  - Eighth-order Butterworth lowpass filter with the cutoff frequency of 90Hz
  - Notch filter with the stop frequency of 60Hz
- Normalized by dividing the signal by its maximum value
- $y_3(n)$ is obtained by passing $y_2(n)$ through a 8-point MA filter



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

The Pan–Tompkins algorithm

- It is a real-time QRS detection algorithm based on analysis of the **slope**, **amplitude**, and **width** of QRS complexes.
- The algorithm includes:

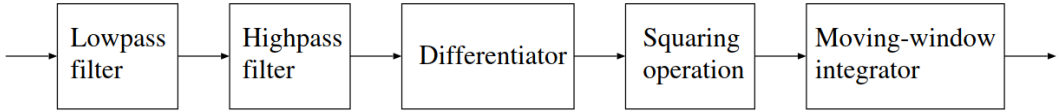


Pan–Tompkins algorithm steps (1)

1. **Lowpass filter**: to reduce computational complexity with a cutoff frequency of 11 Hz
2. **Highpass filter**: with a cutoff frequency of 5 Hz
3. **Differentiator**: approximate the ideal derivative operator ($\frac{d}{dt}$)

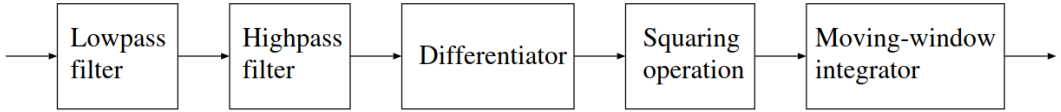
$$y_0(n) = \frac{1}{8}[2x(n) + x(n\ -\ 1) - x(n\ -3) - 2x(n-4)]$$

The derivative procedure suppresses the low-frequency components of the P and T waves, and provides a large gain to the high-frequency components arising from the high slopes of the QRS complex.



Pan–Tompkins algorithm steps (2)

4. **Squaring operation**: makes the result positive and emphasizes large differences resulting from QRS complexes
5. **Moving-window integrator**: after the Differentiator, there are peaks within the duration of a single QRS complex => Moving-window integrator performs smoothing of the output
   - The choice of the window width affect the output:
     - Too small: multiple peaks for a single QRS
     - Too large: the output include the QRS and T waves
     - A window of 30 samples for 200 Hz sampling rate

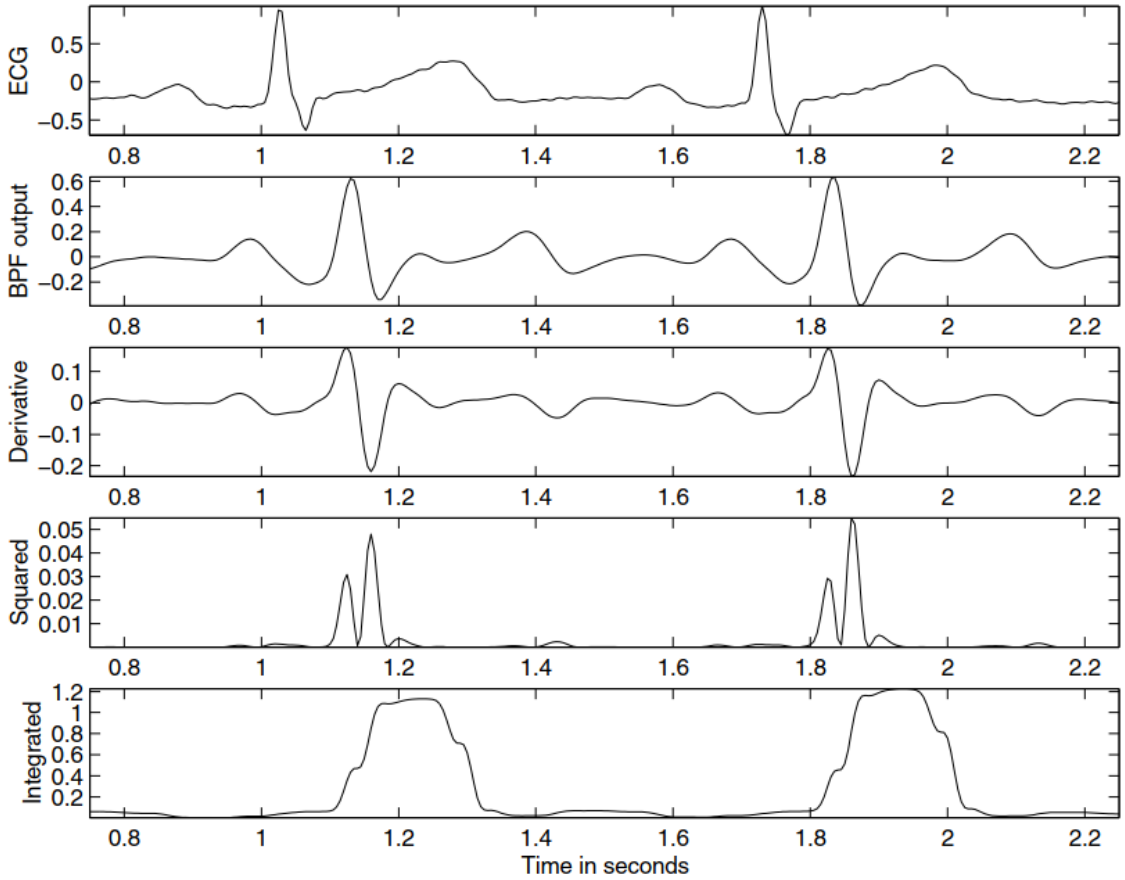


Pan–Tompkins algorithm – adaptive thresholding

- The thresholding procedure changes by computing running estimates of signal and noise peaks
- Two steps are used to categorize peaks detected as **signal (QRS)** or **noise**
- QRS is detected if:
    1. A peak is more than a threshold value. The threshold is updated using previous detected peaks (QRS and noise).
    2. RR intervals of previous peaks (average value) and limits (max and min) of the RR intervals are used to confirm the detected peak is QRS.

Example 1

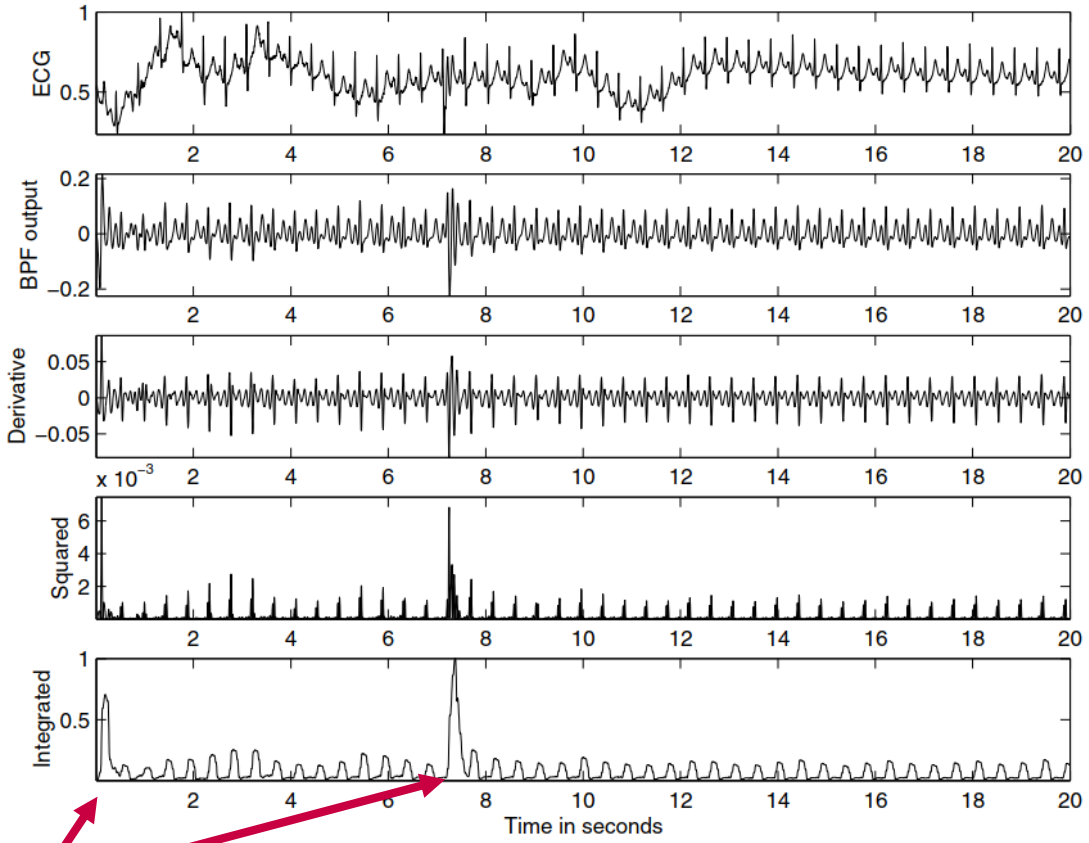
- The derivative operator suppresses the P and T waves
- Observe the shift between the actual QRS location and the pulse output due to the cumulative delays of the various filters.



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

Example 2

- There are two artifacts in the output of the algorithm
  1. Filtering artifact
  2. Artifact in the signal



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

The accuracy of QRS detection

- The Balda et al. (i.e., derivative-based) method is sensitive to noise
- The Pan–Tompkins algorithm[1] had high accuracy
  - 99.3% accuracy in MIT/BIH database
- The algorithm was updated and tested in different setups
- QRS detectors are based on the ones in Pan–Tompkins algorithm with some changes (e.g., different Moving-window integrator)
- Examples:
  - Hamilton Tompkins algorithm[2]
    - 99.69% sensitivity in MIT/BIH database
  - Hamilton algorithm[3]
    - 99.8% sensitivity in MIT/BIH database

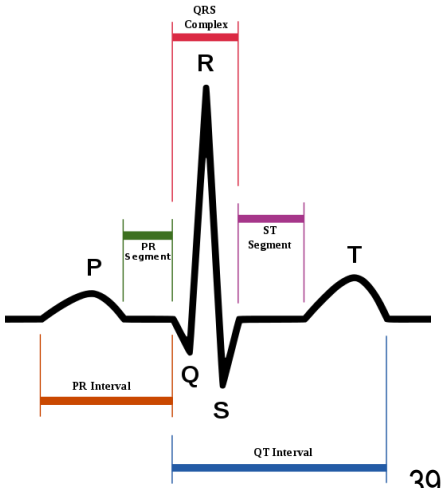
1. Pan, J., & Tompkins, W. J. (1985). A real-time QRS detection algorithm. IEEE Trans. Biomed. Eng, 32(3), 230-236.
2. Hamilton, P. S., & Tompkins, W. J. (1986). Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE transactions on biomedical engineering, (12), 1157-1165.
3. Hamilton, P. (2002, September). Open source ECG analysis. In Computers in cardiology (pp. 101-104). IEEE.

ECG cycle (1)

a) The R-peak in QRS complex is detected

b) The boundaries of the QRS complex is extracted:

- Nearest zero-crossing points
- The time distance should be in the limits of the wave duration
- The maximum slope in the detected wave should be bigger than a threshold

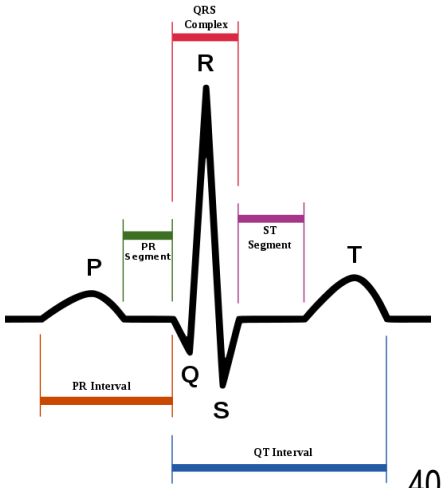


ECG cycle (2)

c) The QRS complex is replaced with the baseline (a few samples preceding)

d) T wave and P wave are detected
- There are different methods proposed [1] [2]
- Detection of P wave is more difficult as it has lower amplitude
- The peaks are detected and then the lengths are extracted

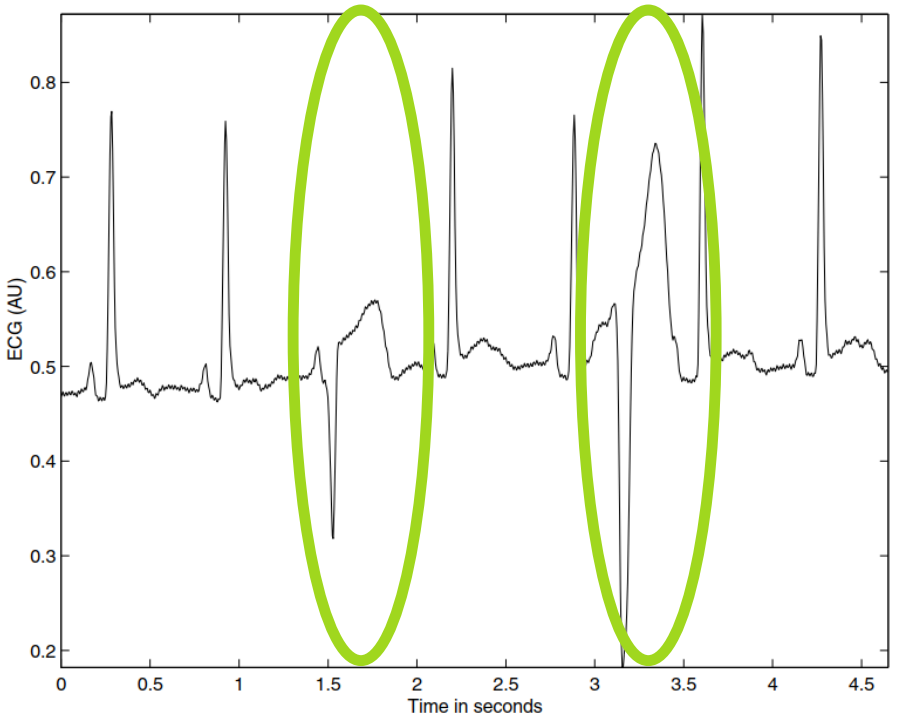
e) Other segments are extracted:
- ST segment
- PR segment



1. Laguna, P., Jané, R., & Caminal, P. (1994). Automatic detection of wave boundaries in multilead ECG signals: Validation with the CSE database. Computers and biomedical research, 27(1), 45-60.
2. Hengeveld, S. J., & Van Bemmel, J. H. (1976). Computer detection of P-waves. Computers and Biomedical research, 9(2), 125-132.

Notes

- ECG segmentation is important as arrhythmias affect the shape of the components
- These components can help us to detect abnormalities in the biosignals



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

ECG feature extraction

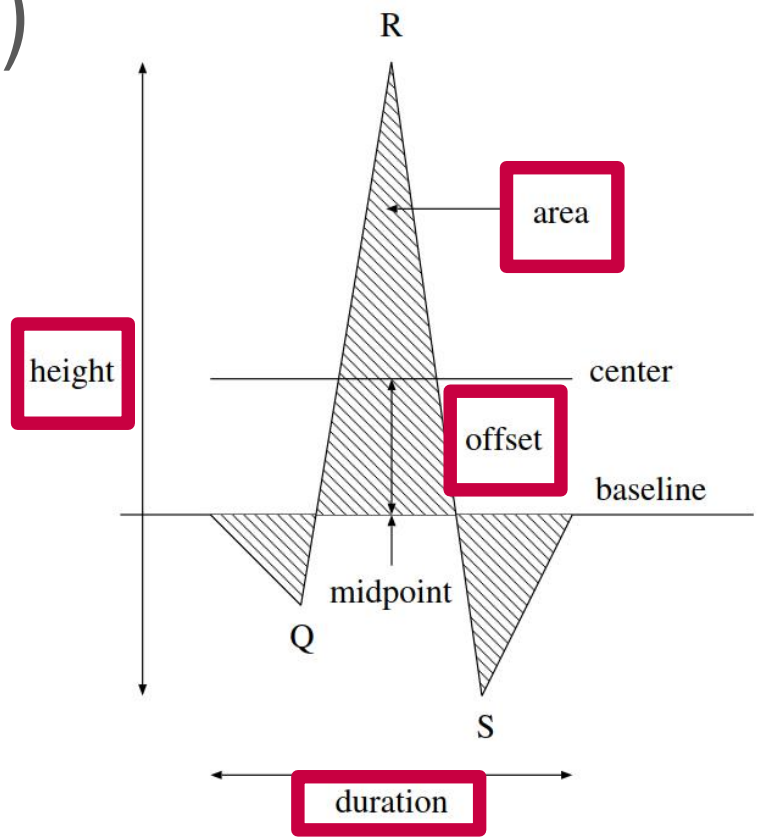
- Different features can be extracted to differentiate normal biosignals from abnormal biosignals
- The features show the characteristics of the events (signals)
- The levels of feature extraction and analysis are determined according to the application
  - R peak detection is enough for heart rate detection

Time-domain features

- Some examples of time-domain features
  - Height
  - Offset
  - Area
  - Duration
  - Interval between two successive components

QRS complex features (1)

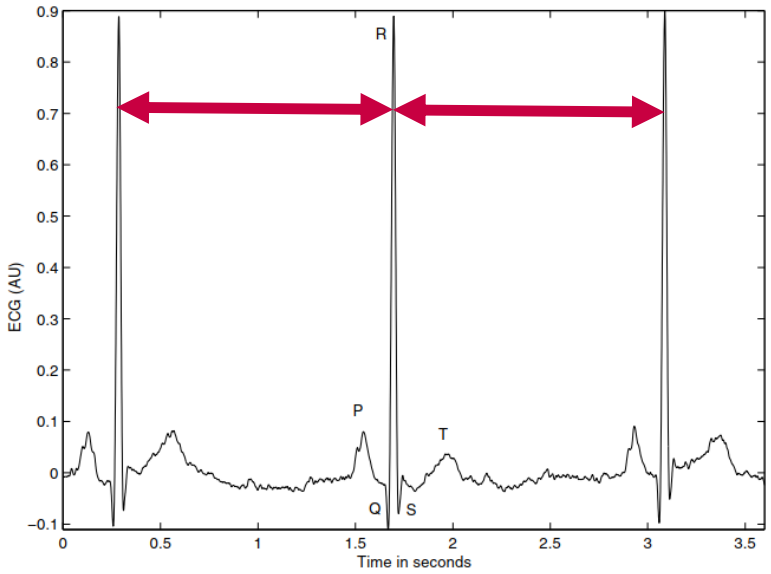
- **Duration**: the duration or width of the QRS complex
- **Height**: the maximum amplitude minus the minimum amplitude of the QRS complex
- **Offset**: the vertical distance from the midpoint of the baseline to the center of the QRS complex.
- **Area**: the area under the QRS waveform rectified with respect to a straight line through the midpoint of the baseline.



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

QRS complex features (2)

- **Interval:** interwave intervals between two successive components
  - R-R intervals
- New values are related to preceding intervals



Rangayyan, R. M. *Biomedical signal analysis*. 2nd Edition, Vol. 33. John Wiley & Sons, 2015.

R-R intervals

- **Hear rate (HR)**:
  - $HR = \frac{60}{RR_{interval}}$
  - Or Number of RR peaks in a minute
- **Heart rate variability (HRV)**: variations in the heart cycles (or R-R intervals)
  - Can be affected by different factors
    - E.g., stress and physical activity.
  - Can be used for stress assessment

Conclusion

In this session, we learned about:

- ❑ Correlation techniques in biosignals (e.g., EEG)
- ❑ Waveform analysis of ECG

In the next session, we will learn about :

- ❑ Frequency domain analysis (e.g., PPG)
- ❑ Waveform analysis of EMG and PCG

Thank You

Questions?

Turun yliopisto
University of Turku