DSAA Major Project

Team Members

Neal Karpe - 20161159
Gunjan Karamchandani - 20161086
Jayitha .C - 20171401

Problem Description

Fitness tracking is gaining massive popularity due to the advent of wearable devices that can track your vital signs. Heart rate monitoring is one such feature in many devices such as smart-watches and wristbands. The heart rate is estimated in real time and can guide. People who exercise to adjust their workload and training programs, which is especially useful in rehabilitation. Most of the recording is done using photoplethysmographic (PPG) signals which are recorded from the wearer's wrist. The PPG signal is recorded using embedded pulse oximeters. A pulse oximeter records a signal by illuminating the skin with an LED and measuring the intensity changes as the light reflects off the exercises during the wearer's skin, forming a PPG signal. Each cycle of the PPG signal corresponds to a cardiac cycle, thus the heart rate can be estimated from the periodicity of the PPG signal.

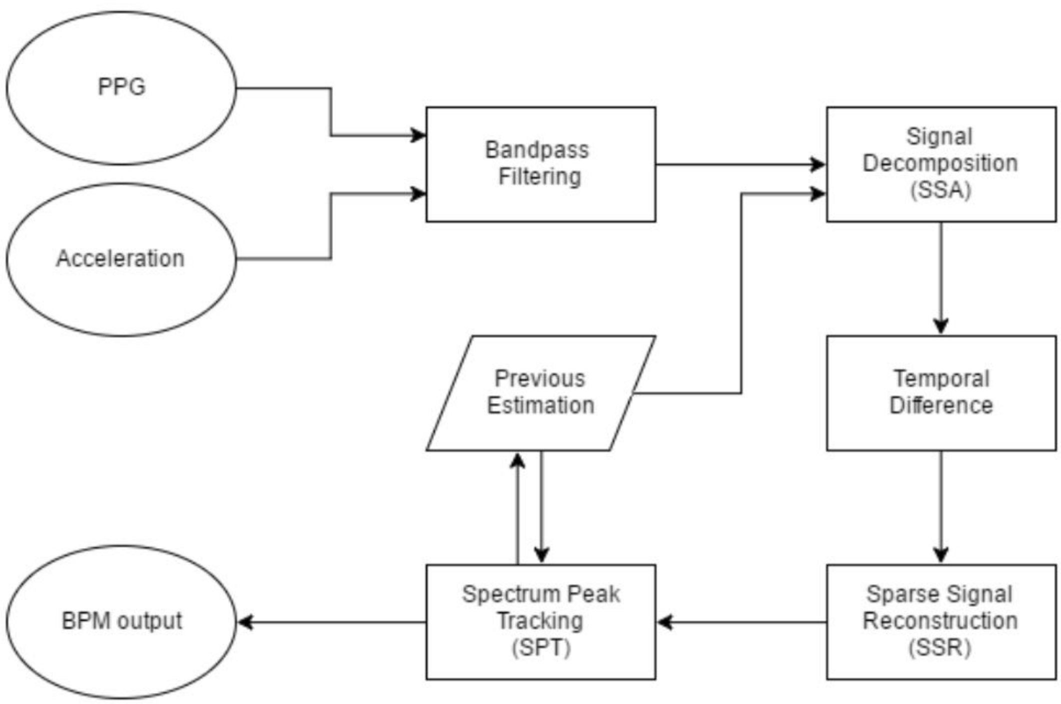
Problem Statement

Data may have noise because of excessive motion during exercise and loose contact between skin and wearable device. Therefore the data provided in the PPG signal isn't sufficient to calculate the BPM. The objective is to come up with an algorithm and implement it to give the estimated heart rate in each time window of 8 seconds where two successive time windows overlap by 6 seconds while eliminating the influence of motion artifacts caused due to excessive relative movement between the device and the wearer

Solution Proposed - Using TROIKA Framework

Briefly

TROIKA is a framework that consists of three key parts: signal decomposition, Sparse Signal Reconstruction (SSR) and Spectral Peak tracking (SPT). Before the PPG signal goes into the signal decomposition state, it goes through a bandpass filter. This filter removes frequencies that cannot be associated with a humanly possible heart beats. The signal decomposition stage is used to denoise the PPG signal and remove motion artifacts. this is done by decomposing a single signal into multiple components. The components can then be analysed and the ones associated with noise and interference can be removed. After that the signal is reconstructed without those components. Before the signal goes through the SSR stage, it is temporally differentiated to remove MA components that correspond with aperiodic movements. The SSR stage gives a high resolution spectrum estimation of the signal, which makes the peak tracking easier. The SPT stage analyses the signal that is created by SSR and selects the spectral peaks that correspond with the heart beat. It makes use of the frequency harmonic relation of heart rates and that the heart rate can't make big changes in two successive time windows.



Signal Decomposition Algorithm - Singular Spectrum Analysis (SSA)

Why SSA?

SSA is a powerful technique of time series analysis. It can be used in with many applications and works with arbitrary statistical processes. This includes linear and non-linear signals, stationary and non-stationary signals and so on. SSA decomposes a time series into oscillatory components and noise. A disadvantage of the SSA algorithm is that it does not automatically remove MA. This is because there is not enough information to determine which components contain the heart rate signal. However there is a solution to this: accelerometer data can be used to recognize MA components in the SSA decomposition and they can be excluded from the reconstruction.

SSA - Algorithm

The steps for decomposition are:

1. Computing the L-trajectory matrix X . This transfers a one-dimensional time series $Yn = (y1, \ldots, yn)$ into the multidimensional series $X1, \ldots, Xk$ with vectors $Xi = (yi, \ldots, yL)^T$. L is the window length and is chosen such that 2 ≤ L ≤ N. K is chosen such that K = N − L + 1. The result is the trajectory matrix X = $[X1, \ldots, Xk]$

2. Compute the matrix $S = XX^T$. This is a positive definite and symmetric matrix which can be decomposed using eigenvalue decomposition. Also the eigenvalues of this matrix are real and positive.

3. Singular value decomposition (SVD) of the matrix S: Compute the eigenvalues and eigenvectors of the matrix S. If λ1,..., λL are the eigenvalues of S in decreasing order and U1,..., UL are the corresponding eigenvectors of these eigenvalues, then $Vi = X^T\ Ui/\sqrt{\lambda i}$ (i = 1; : : : ; d). Here d = max{i; such that λi > 0}. The trajectory matrix X can now be written as:

   $X = X1 + \ldots + Xd$

   where Xi = $\sqrt{\lambda i}\, UiV_i^T$. The collection ($\sqrt{\lambda i}$; Ui; Vi) is called the eigentriple of the SVD.

The steps for Reconstruction are:

1. Grouping: The elementary matrices Xi are split into m disjoint subsets I1; : : : ; Im. Let I = i1; : : : ; ip, then the matrix XI corresponding to the group I is defined as XI = XI(1) + : : : + XI(p). The trajectory matrix X can now be written as: X = XI(1) + : : : + XI(m)

2. Diagonal averaging: Each matrix XI is transformed into a new time series of length N. These are the components that the original time series is split into. Let X be an L × K matrix with elements $X_{i,j,}$ 1 ≤ i ≤ L, 1 ≤ j ≤ K. Then L^* = min(L; K); K^* = max(L; K) and N = L + K − 1. Let x^*_{ij} = xij if L < K and x^* ij = xji otherwise. Then diagonal averaging transfers the matrix X to the series g0; : : : ; gN−1 by the following formula:

$$g_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} x^*_{m,k-m+2} & \text{for } 0 \leq \text{k} < L^*\text{- 1}, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} x^*_{m,k-m+2} & \text{for } L^* \text{ - 1} \leq \text{k} < K^*, \\ \frac{1}{N-k} \sum_{m=k-K^*+2}^{N-K^*+1} x^*_{m,k-m+2} & \text{for } K^* \leq \text{k} < \text{N}. \end{cases}$$

Removal of Motion Artifacts - Removal of Components (RoC)

Briefly

The objective is to identify the components that correspond to MA and reconstruct the signal without those components. The main idea is that the components that have a peak at approximately the same frequency as the peaks of the acceleration data (MA components) in the frequency domain are identified so that they can be excluded from the reconstruction.

Identification of MA components

Kurtosis of the signal is used to check whether the signal contains noise or not. The kurtosis of a signal can tell how peaked the signal is. A low kurtosis means that the signal is flat, which means that it consists mainly out of noise. The next step is to locate the peaks of the acceleration data that are responsible for the MA peaks in the PPG signal. When the amplitude of the peak in the acceleration data is small, the corresponding peak in the PPG signal will be small too. For this reason only the peaks that have an amplitude larger than 50% of the maximum peak in the acceleration data are considered.

The following step is to compare the locations of those peaks to the location of the maximum peak of each component of the PPG signal that is decomposed by SSA. Since the peaks of the acceleration data do not always exactly coincide with the MA peaks in the PPG signal, some tolerance (MA tolerance) has to be set.

Components that have their maximum peak at around the same frequency as one of the peaks of the acceleration data, will not be removed if they are close to the frequency of the previous estimated HR. However the HR can change significantly between two time windows. This means that here too some kind of tolerance has to be set, the BPM tolerance.

Spectral Peak Tracking (SPT)

SPT of RoC also considers the peaks outside the BPM-interval, in some cases. If the SPT of the RoC can not find a dominant peak three times in a row within the BPM tolerance interval, it will search the whole spectrum for the maximum peak. This is the retracking method for the SPT of RoC.

Temporal Difference

Temporal difference is the name given to the operation that returns the difference between sequential values of a signal. If y = [h(1), h(2),...,h(M)], then the first order difference is defined as y′ = [h(2) − h(1), h(3) − h(2), ..., h(M) − h(M − 1)]. The second order difference for y is the first order difference of y′. As long as k is not large, the k-th order difference maintains the fundamental and harmonic frequencies of the signal. This means that non-periodic frequencies will be removed. Because the PPG component that is associated with the HR is approximately periodic in a short time window and MA is generally aperiodic, the temporal difference is applicable here.

Sparse Spectral Reconstruction (SSR)

This is a pretty standard algorithm that searches for the HR peak using previously estimated HR values

Optimising Parameters

- L - Generally for the window length L it is good to choose a value that is less than half the length of the signal that is going to be decomposed, so that the error can be balanced and there is the ability to resolve lower frequencies.

- In the grouping stage the matrices that belong to the eigentriples where the eigenvalues are close to each other in size are grouped together. This is done because those matrices are basically equal to each other. Eigentriples that have an eigenvalue that is close to zero, contain no significant information so they can be excluded from the reconstruction.

- By analysing the data, It is found when there are clearly distinguishable peaks in the acceleration data, the kurtosis of the signal is above the value of 120.

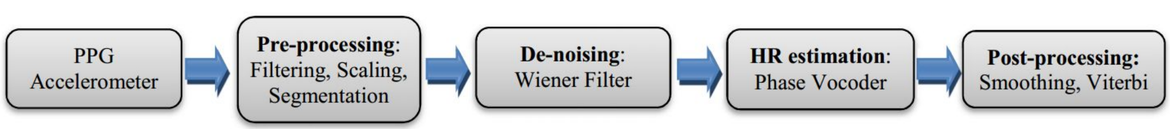
- MA Tolerance - This tolerance has to be large enough that it includes the components that correspond with MA but small enough that it does not include the true HR when the MA is close to that frequency. When analysing the data sets, it is seen that the maximum difference between the locations of the peaks in the acceleration data and the corresponding peaks in the PPG signal is around 5 BPM. Thus a tolerance of ±5 BPM should be sufficient to determine the components that belong to MA

- BPM Tolerance - The maximum change in the HR between two successive windows can be determined. From this it is given that the HR can change by a maximum of 10 BPM..

Alternative Solution proposed - Using Wiener Filter and Viterbi Decoding (WFPV+VD)

Briefly

The algorithm exploits a Wiener filter to attenuate the motion artifacts, a phase vocoder to refine the HR estimate and user-adaptive post processing to track the subject physiology. Additionally, an offline version of the HR estimation algorithm that uses Viterbi decoding is designed for scenarios that do not require online HR monitoring



Preprocessing

1. The two PPG signals and three accelerometer signals are filtered with a 4th order Butterworth band-pass filter (0.4-4 Hz)
2. The two PPG signals are then normalized to zero mean and unit variance and averaged

3. The averaged PPG signal and the 3 accelerometer signals are down-sampled from 125 to 25Hz
4. 1024-point DFT is performed on the signals

De-noising using Wiener filtering

Why Wiener Filtering ?

The Wiener filter is a common tool to estimate a desired signal by linear time-invariant filtering of an observed noisy process. Assuming known stationary signal and additive noise spectra, the Wiener filter performs the minimum mean square error estimation of the desired signal given another related process.

How is it used ?

The noisy PPG signal, X(f), is assumed to be corrupted by additive MA noise:

$$X(f) = S(f) + N(f)$$

where $S(f)$ and $N(f)$ are the spectra of the clean PPG signal and the MAs, respectively.

The estimation of the clean signal can then be obtained as:

$$\widehat{S}(f) = X(f) - N(f) = \left(1 - \frac{N(f)}{X(f)}\right) X(f) = W(f)X(f)$$

For a signal observed in uncorrelated additive random noise, the frequency-domain Wiener filter is given as:

$$W(f) = \frac{Pxx(f) - Pnn(f)}{Pxx(f)} = \frac{Pss(f)}{Pss(f) + Pnn(f)}$$

where $Pss(f)$, $Pnn(f)$ and $Pxx(f)$ are the power spectrums of the clean signal, noise and observed signal.

Principle behind using Wiener Filter - Convolution Theorem

The filter convolution in time domain is equivalent to multiplication in frequency and thus the Wiener filter acts as an adaptive signal-to-noise dependent attenuator, where frequencies which are more affected by the noise are given less importance.

Estimating Power Spectrums

The noise spectrum can be directly estimated from the accelerometer signals which is done by averaging the spectrum of the 3 accelerometer signals. The clean PPG spectrum, $Pss(f)$, can be estimated as a subtraction of the noise signal from the observed signal, $Pxx(f) - Pnn(f)$, or recursively from previous filter outputs.

How are the Wiener Filters Implemented ?

Depending on how the power spectrum of the clean PPG signal is estimated, two Wiener filters are implemented, with frequency domain filter coefficients given as:

$$Wa(t,k) = 1 - \frac{Pnn(t,k)}{\frac{1}{C}\sum_{i=t-C+1}^{t} Pxx(i,k)}$$

$$Wb(t,k) = \frac{\sum_{i=t-C}^{t-1} Wb(i,k)Pxx(i,k)}{\sum_{t=t-C}^{t-1} Wb(i,k)Pxx(i,k) + C.Pnn(t,k)}$$

where w(t,k) is the weight of the k-th frequency bin at time, t.

HR Estimation and Refinement using Phase Vocoder

Why Phase Vocoder?

The phase vocoder technique is employed to refine the initial HR estimate through the estimation of the instantaneous frequency as the rate of change of phase angle at time. The effective frequency resolution (the minimum frequency that can be estimated, the Rayleigh frequency) of the data is limited by the size of the window of the analyzed data (8s) and equals to 1/8*60 = 7.5 BPM. Zero-padding before DFT is used to interpolate the spectral envelope to other frequencies thus decreasing the frequency spacing between neighboring DFT bins. This does not create new information but allows for a better revelation of the existing information in the signal. The phase vocoder is the technique that is used in audio processing to manipulate audio length without changing its pitch or to change its pitch without affecting its length, by preserving the coherence of phase information. The phase vocoder uses a polar representation of the DFT and the instantaneous frequency estimation is computed as a discrete derivative of the phase. When analyzing the signal with multiple overlapping windows the individual signal components (sinusoids) will be correlated in time and spread over multiple adjacent DFT frequency bins (spectral leakage).

Instantaneous Frequency $$f(t) = \frac{1}{2\pi}\frac{d\theta(t)}{dt}$$

The DFT phases, θ2, θ1, from the current and previous frames, of the chosen frequency peak in the magnitude spectrum, , are used to refine the initial frequency estimation:

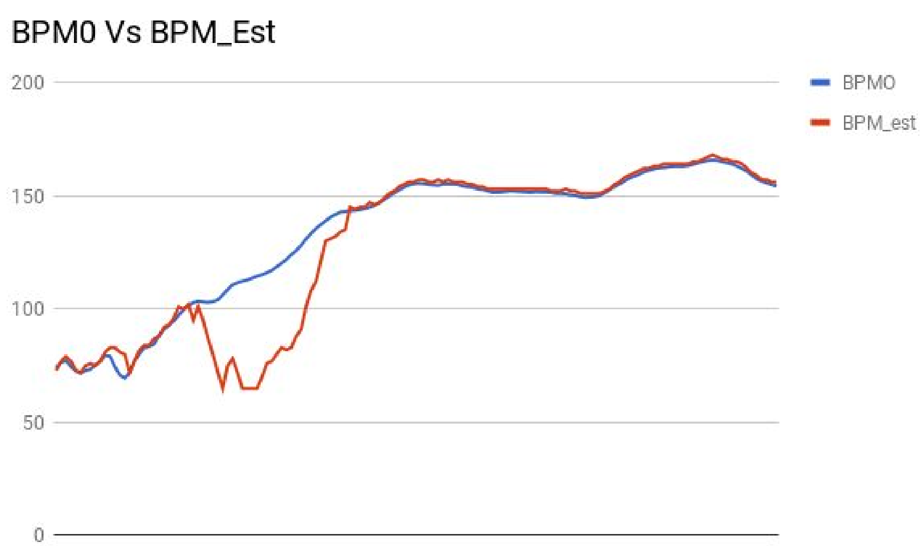
$$\arg\min_{n}\left(\tilde{f}(n) - f\right);\ \tilde{f}(n) = \frac{(\theta_2 - \theta_1 + 2\pi n)}{(2\pi(t_2 - t_1))}, \forall n \in \mathbf{N}$$

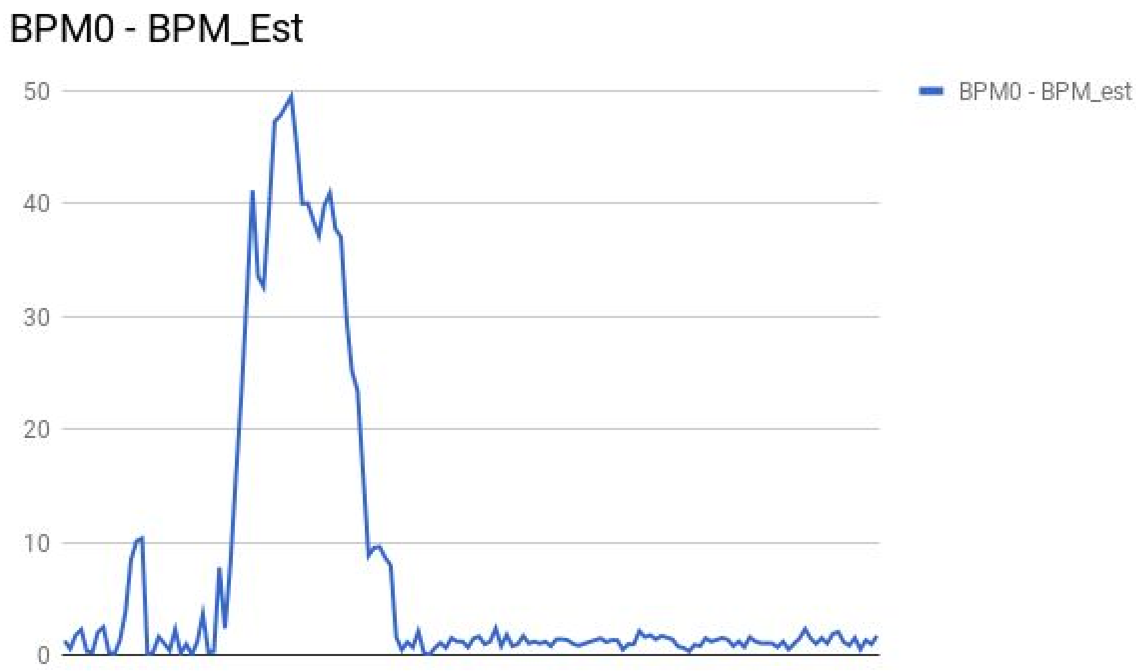
Post Processing - Viterbi Decoding (VD)

he offline version of WFPV substitutes the original post processing steps with a probabilistic framework using Viterbi decoding. The time-frequency plane (spectrogram) of a complete recording which is composed of DFT magnitudes after Wiener filtering is considered as a N-by-T state-space map of emission probabilities, B, for N states (discrete values of HR) and T observations (time windows), where bjt is a magnitude value of the jth DFT bin for the tth time window. The N-by-N matrix of transition probabilities, A, where aij represents the probability of changing from the ith HR to the jth HR, is estimated from the ground truth automatically by counting the transitions using the leave-one-recording-out procedure. In this manner, the ground truth of the testing recording is never used but the ground truths of all other recordings are used to estimate the transition probability matrix.

Analysis of the solution

Analysis done on one particular training set





The average error was found to be ~ 7.411 BPM

How to Improve our Solution?

TROIKA uses Signal Decomposition techniques to remove MAs, but using Spectral Subtraction is a better alternative. Here the acceleration spectrum is subtracted from the PPG spectrum. This removes peaks corresponding to MA from the PPG spectrum. The result of this is a cleansed spectrum used by SPT to produce a BPM output. This output is can then be used as feedback for SPT to search in smaller intervals.