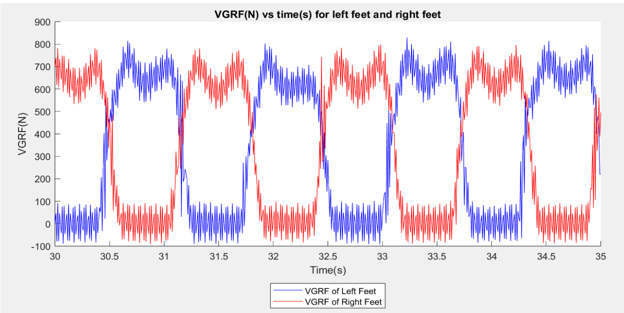
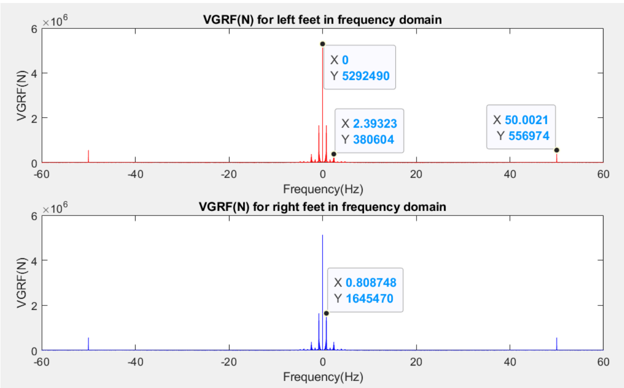
Introduction

The walking pattern of a human is known as gait. A gait cycle, or simply a stride, can be divided into two major phases, namely stance and swing. The aim of this project is to extract the sequence of stance times, swing times and stride times for each foot from the given noisy VGRF data. Data as such could be useful in improving healthcare.

Data Exploration

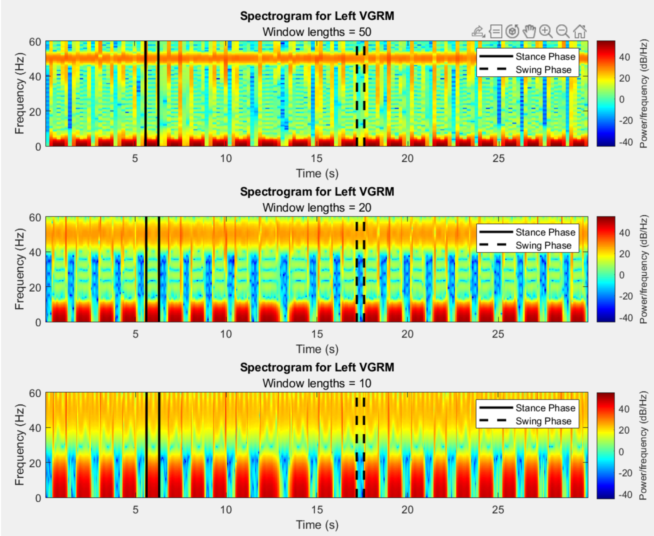
3.1 Initial plots





The plot on top shows the time domain representation of raw VGRF data for both feet. At the start of the plot, it can be seen that when the right foot is experiencing high VGRF (Vertical Ground Reaction Force), left foot experience near zero VGRF, and after approximately 0.8 seconds, the VGRF under each foot becomes the opposite. It is also observed that this pattern repeats itself every 1.3 seconds. This suggests that the gait cycle (stance phase followed by swing phase) of this subject lasts for roughly 1.3 seconds. High frequency noise is also very apparent in this set of data. As for the plot below, which is VGRF data of both feet represented in the frequency domain, a few spikes are evident: the highest spike, appearing at zero frequency, could be due to normal reaction force that always acting whenever feet were on the ground, 0.8 Hz component could possibly be the frequency of gait cycle, 2.3 Hz component could be the frequency of VGRF change during stance, with the local minima (referring to time domain plot) being midstance, and the two peaks correspond to heel strike and toe off. The highest frequency spike, located at around 50 Hz is most likely sensor noise. Upon zooming in, there are also small amounts of different frequencies scattered within 10 Hz, which may be the other frequency component of VGRF.

Answer to Section 3.1 Question: Referring to the time domain plot of VGRF, a simple threshold value, i.e. 120N can be estimated by directly inspecting the plot. Every value above the threshold is regarded as foot on the ground (stance) while values below threshold indicate foot off the ground (swing). However this would sacrifice accuracy due to the presence of noise. Ideally, swing occurs when VGRF is 0 N, but if threshold value of 120N is used (to account for noise), heel strike and toe off could possibly be considered as swing phase. Simple threshold method is not suitable for acquiring accurate stance or swing time without cleaning and filtering the data beforehand. As a side note, this method is not applicable in the frequency domain.

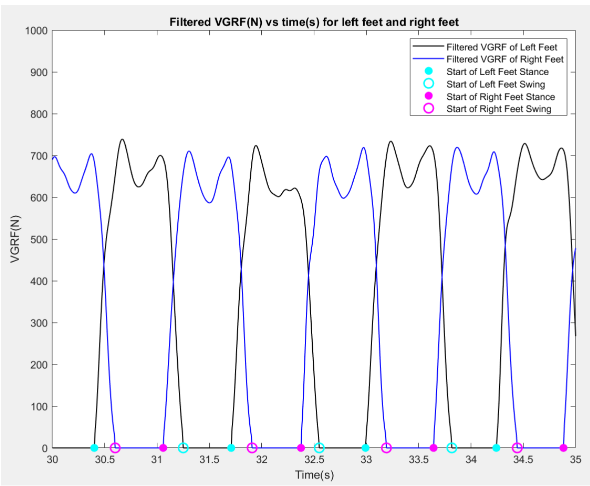


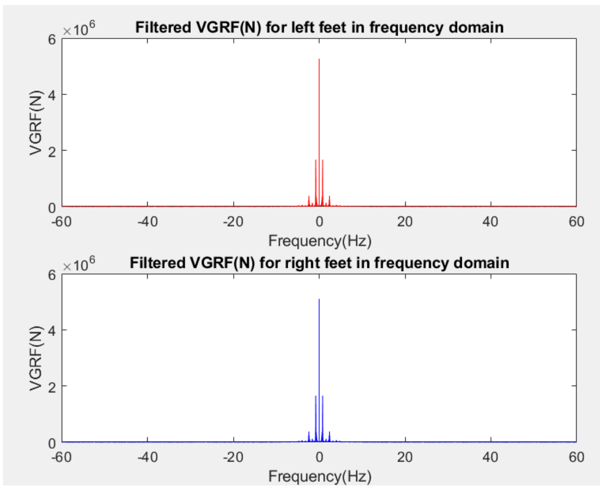
3.2 Time-Frequency Analysis

As shown on the left, spectrogram has allowed time-frequency analysis, where magnitude of each frequency content at each point in time can be visualised through the colour. Warmer tone corresponds to greater magnitude of the signal in that frequency, while cooler tone indicates lower magnitude. As stance phases exert much higher force on the ground, the red colour regions just above the x-axis represent stance phases while the region in between are swing phases. Some stance and swing phases are also marked in the spectrograms.

Answer to Section 3.2 Question: The spectrograms are plotted in the order of descending window lengths. As the window length decreases, it can be observed that spectrograms become less pixelated and smoother. The changes in power with respect to time is more obvious and easier to visualise. However, from the perspective of power against frequency, the power becomes less accurate as window length increases. This is deduced by inspecting all the major frequency components of the signal in Section 3.1 frequency domain plot. One possible cause to this issue is that decreasing window lengths allow changes to be smoother in the time domain, but the window is too short to have enough cycles of periodic signal for Short Time Fourier Transform to output the correct frequency. The main noise component would be the sensor noise mentioned in the previous section, located at around 50 Hz.

Filtering

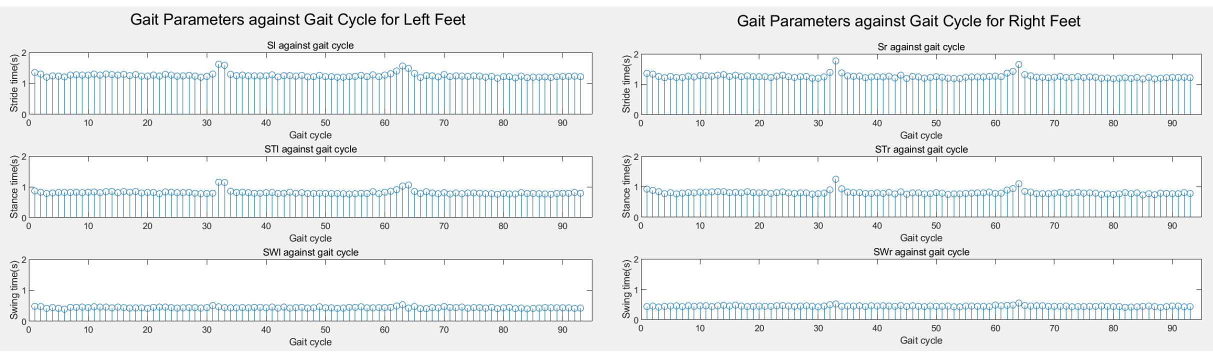




The left plot shows filtered VGRF signal against time while the right plot shows frequency representation of filtered VGRF signal for both feet. To remove the 50 Hz sensor noise, a low pass filter can be used to only allow frequency components lower than a certain threshold value to remain, in this case, the threshold value would be 10Hz. This threshold value is inferred from the observation made in Section 3.1 frequency domain plot, where all VGRF related frequencies are within the range of 0 Hz to around 10 Hz. This is also evident in Section 3.2 spectrogram with window length of 50 (this spectrogram shows the most accurate frequency among all three spectrograms), in which the colour-alternating region near 0 Hz that highly resembles VGRF in time domain plot does not exceed 10 Hz. While there could be other sources of noise within this range, the effect of these noises are negligible in this project, and this is shown in the filtered signal time domain plot, where little to no unexpected fluctuation appears in the VGRF signal. Matlab function `firpm(n,f,a)` is used to filter the noise, where the input to the function are filter order, normalised frequency point with the unit of π/rad, and desired amplitude respectively. Normalised frequency point vector is chosen to be `[0,0.1667,0.2,1]`, and desired amplitude being `[1,1,0,0]` as low pass filter is to be implemented. The corresponding frequency in Hz for the normalised frequency point would be 0 Hz, 10 Hz, 12 Hz and 60 Hz, with 10 Hz being the cutoff frequency, and 10 Hz to 12 Hz acting as a transition band. Filter order of 500 was first chosen as benchmark, then slowly reducing the filter order, comparing the resulting gait parameters stem plot (which will be discussed in the next section) of current filter order with benchmark result, if there are noticeable changes or errors, the value before it is chosen. The rationale behind this is to reduce computing time while not compromising accuracy of output. In the end, a filter order of 150 is selected. To account for the delay of $filter\ order/2$ induced by convoluting the VGRF signal and filter, the resulting signal is truncated to obtain the original signal length.

Answer to Section 3.3 Question: Comparing the filtered VGRF time domain plot with raw signal plot, it is very evident that high frequency fluctuation in VGRF is not present anymore, leaving only clean lines. There is no negative VGRF value as well in the filtered plot. As for the filtered frequency domain plot, there are no major frequency components ranging from 10 Hz to 60 Hz, unlike frequency domain plot from Section 3.1 where there is spike located at around 50 Hz , which means that the filter had greatly attenuated the spike.

Computing Gait Parameters



The left plot shows three subplots of gait parameters against gait cycle for left foot while the right plot shows the same subplots but for right foot. The three gait parameters are stride time, stance time and swing time.

After performing filtering on data as discussed in the last section, a relatively clean signal is obtained. However, there are still small spikes observed in some of the swing phases. Therefore, simple threshold value method is implemented to remove these spikes. This threshold method is exactly the same as what was explored in the previous section, but a lower threshold value is used instead. By using a lower threshold value, the function is able to yield more accurate results, as toes off and heel strike in stance phase are less likely to be categorised as swing phases. In this function, a threshold value of 30 is chosen, thus any VGRF below 30 N will be set to 0 N, implying that these values of VGRF are considered as swing phases. After this, all non-zero VGRF values would indicate stance phases while zero valued VGRF represent swing phases.

The next step would be to extract swing phase and stance phase onset time from the cleaned VGRF signal. Since the end of stance phase is always followed by the start of swing phase, and similarly, the end of swing phase will definitely be followed by the start of stance phase, knowing the start and end indices of either phase would allow indices of another phase to be found as well. In this case, Matlab function `islocalmin(A,'FlatSelection','first')` is called to determine the start of swing phase and `islocalmin(A,'FlatSelection','end')` is used to find the end of swing phase, with an extra condition of VGRF being zero. However, if the last phase is a swing phase, then the function will not work as intended, thus manual search using a loop is required. Then, two logical arrays of the length of VGRF raw data are returned (these logical arrays will be referred to as start and end indices array of swing phase). The start indices array are then shifted to the left to obtain end indices of stance phase, while the end indices array are shifted to the right to acquire start indices of stance phase, empty elements of these new logical arrays are filled in with false. After that, the first element and last element of VGRF data will be categorised as the start or end of one of the phases using the if-else conditional statements.

After finding all the indices, duration of stride, stances and swings can be determined by subtracting end time by start time, and adding $1/Fs$ second to take into account the time of the first index. To ensure the accuracy of data, incomplete gait cycles will be removed. For example, the first ‘cycle’ starts from midstance instead of heel strike, and is therefore removed. Last ‘cycle’ will also not be considered as it is certainly incomplete as well. If the last ‘cycle’ ends with a swing phase, but does not proceed to heel strike (stance), the swing phase is still considered incomplete.

These steps are then repeated for VGRF data of another foot.