

# Picram dataset

Nazrul Anuar Nayan

## I. OBJECTIVE

- 1) To update the status of PICRAM and CALMS-2 datasets.

## II. COMPLETED TASKS

### A. Data structure

1) *PICRAM*: The main purpose of this work is to ensure that all the .csv data are converted into .mat file, without leaving any data out before the respiratory rate estimation process starts. To realize this, several problems have been encountered. Firstly, the size of the data is big. In ProgReport14, there are some ECG data file which have more than 2 Gb and was unable to be read using 'xlsread'. To solve this, this week, 'textscan' have been utilized after reviewing some comments in the Matlab forum. After specifying the number of columns, textscan is set to read 10 million (mil) rows. Then, the data is saved as ECG1 (or PPG1) into bspprojects9\OURR\_results\Picram. The data are then removed from Matlab workspace. It continues to read the next 10 mil rows and so on. This reduces the usage of memory by 33%. As an example, ECG for Oxford patient '012' contains 2.5 Gb data. There are 90,548,950 rows of data in the file. These ECG data has been divided into 20 mil rows each. For this method, using the current codes, it takes 9 hours to read all those 90 mil rows. Eventhough the process can still be improved to shorten the time using some other '.csv' read function, it is still practical since the computer did not freeze due to memory insufficient. At this moment, the data are still being processed and uploaded to bspprojects9\OURR\_results\Picram. The final data will be uploaded to bspprojects9\OURR\Picram for respiratory rate estimation use. Table I, shows the latest update on the two datasets.

For this week, I have also downloaded JMP software (trial version) to directly load .csv file and plot the graph to overview the data prior processing. The software can load up to 4 mil rows of 6 columns. The time taken to load are less than 5 minutes. Fig.1 shows the graph plotted using JMP for Oxford 011 patient's PPG.

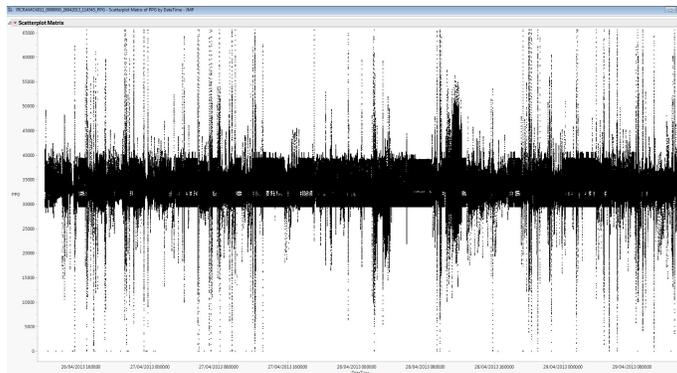


Fig. 1. PPG waveforms from .csv file using JMP

2) *CALMS2*: As shown in Table I, raw CALMS2 dataset are located in bspprojects5\cancer\_hospital\Mat\_files. From the records, there are 347 patient waveforms. 'Notes on BSP Patient Data' created by Marco has been checked. The analysis of the data will start next week.

## III. CONCLUSION

This week, the work to upload the data of PICRAM has been carried out. It has not completed yet, but the method used seems to be working. Overview of the data using JMP has also been done. The work on CALMS-2 data will start next week.

TABLE I  
DATASETS UPDATE

	CALMS2	PICRAM
Data updated	2014	2015
Formatted/ ↑ bspprojects9	✗	✓ and still running
Raw data	bspprojects5	bspprojects12
ECG Signal	✓ (only 25)	ICU ✗, ward ✓
PPG Signal	✓	ICU ✗, ward ✓
Accelerometry Signal	✗	✗
Subject Population	post-surgery	ICU-discharged
Clinical Setting	post-op ward	ICU, ward
Breathing (spontaneous/ ventilated)	s	ICU s,v / ward s,v
No. of records	347	ward (Ox 199, Rb 68)
Neonates (<1 y.o.)	✗	✗
Paediatrics (1-18)	✗	✗
Young Adult (19-40)	✓ (very few)	✓
Adult	✓	✓
Elderly (>70 y.o.)	✓	✓
Unwell?	✓	✓
Chronically Unwell?	✗	✗
Acutely Unwell?	✓	✓
Ambulatory?	ward ✓	✓
Recording Time	1:30 h	ward 1-2 d
ECG/PPG Acq. Equip.	PPG: Nonin bluetooth to PDA	PPG: Nonin 4100
Continous RR Signal	IP	✗
Other Reference RR	nurse	nurse (track & trigger)
Avalability	available	available