Physiological state characterization by clustering heart rate, heart rate variability and movement activity information

Niranjan Bidargaddi[*], Antti Sarela[*] and Ilkka Korhonen[‡]

Abstract— The objective is to identify whether it is possible to discriminate between normal and abnormal physiological state based on heart rate (HR), heart rate variability (HRV) and movement activity information in subjects with cardiovascular complications. HR, HRV and movement information were obtained from cardiac patients over a period of 6 weeks using an ambulatory activity and single lead ECG monitor. By applying k-means clustering on HR, HRV and movement information obtained from cardiac patients, we obtained 3 clusters in inactive state and one cluster in active state. Two clusters in inactive state characterized by - a) high HR and low HRV b) low HRV and low HR, could be inferred as pathological with abnormal autonomic function. Further, activity information was significant in differentiating between the normal cluster found in active and an abnormal cluster found in inactive states, both with low HRV. This indicates that the activity information must be taken into account while interpreting HR and HRV information.

I. INTRODUCTION

Heart rate (HR) and heart rate variability (HRV) analysis is being studied with great interest in clinical and physiological research over the recent years due to its potential association with mortality, autonomic balance, exercise responses, sleep disorders, etc [1][2]. Root mean square of successive RR interval differences (RMSSD), Standard deviation of RR interval values (SDNN), Low frequency power (LF), high frequency power(HF), etc. are some of the indexes to quantify HRV. RMSSD is a time domain measure of HRV, sensitive to high-frequency HR fluctuations in the respiratory frequency range (0.2 - 0.4 Hz), and has been used as a measure of vagal tone. It is defined as SD of differentiated HRV calculated by high pass filtering HRV. High HR should be associated with low RMSSD and vice versa.

In cardiac patients reduced HRV indicates a risk factor for future cardiac events [3][4]. Measures derived from short term ECG recordings have proven to be effective predictors of all cause mortality and sudden cardiac death in CAD patients [5]. A low HRV and/or increased HR at any point of time indicates alterations of autonomic activity characterized by sympathetic nervous system activity and/or vagal withdrawal. Parasympathetic nervous system which regulates the vagal tone or HR is dominant during relaxed state and sleep. Vagal activity decreases during exercise and is eventually absent, while sympathetic activity dominates

[6]. Sympathetic nervous system can also be triggered by anxiety, stress, depression and fatigue [7][8][9][10][11].

Despite its broad applications, HR and HRV analysis is done predominantly using short term recordings. Analyzing long term recordings obtained from a free living environment can provide additional behavioral and environmental information. Further, by collection large samples of free living recordings at different points of time, we can overcome the inherent drawbacks of day-to-day random variations observed in short term HRV [12]. Over recent years, availability of low cost unobtrusive wearable physiological monitoring technologies has enabled long term monitoring of ambulatory physiological data. Assisted by these technologies we can study free living physiological dynamics such as circadian HR variation.

The primary objective of this paper is to cluster HR, HRV and movement activity information into sections of normal and abnormal combinations and interpret the significance and implications of these clusters for cardiac patients. The data used for physiological state characterization is obtained from a free living setting. Since sympathetic nervous system could be triggered due to either physical or mental factors, we incorporate movement state information in this analysis. Secondary objective is to assess the clinical significance of movement activity information in interpreting HR and HRV.

The rest of the paper is organized as follows. In section II we describe the monitoring technology, data collection process and the techniques used for analysis. Results are presented in Section III along with discussions and future works in section IV.

II. METHODS

A. Monitoring device

As part of this study, we used personal monitoring device: Alive Heart Monitor, marketed by *Alive Technologies, Gold Coast, Australia, http://www.alivetec.com/*. It provides accelerometer data for every spatial axis with sampling frequency of 75 Hz and a range of +2.7 to -2.7 g. It also records ECG signals sampled at 300Hz through a single lead. At full charge the monitor can log approximately five days ECG and accelerometer data to an onboard SD memory card.

This device can be mounted at various body locations. In this study the device was mounted on patients waist, since it is close to the center of the mass.

B. Data Collection

We collected activity and ECG data from patients undergoing second phase cardiac rehabilitation at North Lakes (QLD,

[*]Australian E-Health Research Centre, CSIRO ICT Centre, Lvl 20:300 Adelaide street, Brisbane, QLD 4000, Australia

[‡] VTT, ICT for Health VTT, P.O.Box 1300, FIN-33101 Tampere, FINLAND
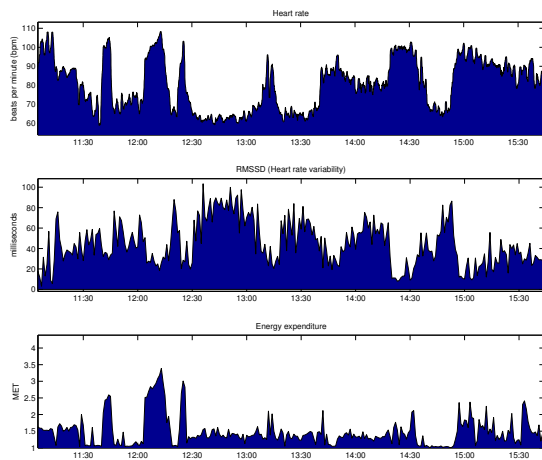
Corresponding author: niranjan.bidargaddi@csiro.au

Fig. 1. Continuous values of HR, HRV and MET from 11:00 to 16:00 of a particular day within six weeks of data collection period for an arbitrarily chosen patient.

Australia),as part of a larger study. In the rehabilitation program, patients attend exercise sessions twice a week at the hospital gym for a duration of six weeks. In each session, patients perform different exercises for a total duration of one hour. Typical exercises are: biking, rowing, walking on stairs, arms exercise, walking on treadmill, and working with exercise ball. At the end of each exercise patients record their pulse rate, SP02 (pulse oximetry), and the Borg scale (exercise difficulty). Each exercise activity lasts from 3 to 7 minutes and the physiotherapist can vary the intensity settings and duration of the exercise to maintain a steady recommended difficulty level (viz., Borg scale = 11).

The patients wear the accelerometer device for entire duration of rehabilitation. However, due to the obtrusive nature of ECG leads, they wear them mainly during exercises and at least once a fortnight all day.

The data logged on the SD memory card is downloaded twice a week by the physiotherapist when patients attend the rehab session for further analysis. The batteries are also changed at the same time. For this study, we created a data set by aggregating the ECG and corresponding activity recordings present in 6 weeks for each of the 7 patients who had undergone surgery due to cardiovascular disorders (Sex: 4 males, 2 female; BMI: 18.85 - 34.28 (median 27.16); Condition: 5 with Coronary artery and 1 with valve diseases; Depression diagnosed: 2; beta blockers: 4; smoking history: 4; Medications: 5-6; Hypertension: 4). The data set for each patient contained predominantly two full day recordings two weeks apart and 6-10 one hour recordings during the exercise sessions.

C. Analysis

We derived HR and HRV from ECG recordings after removing arrhythmias using Firstbeat Pro software, (*Firstbeat Technologies, Inc, Jyvaskyla, Finland, www.firstbeattechnologies.com*). Both HR and HRV were normalized based on maximum and minimum values for each patient, to balance the variations in absolute values between patients due to medication effects and age. We used RMSSD values calculated using one minute time window as a measure of HRV.

In a separate study (unpublished), we derived metabolic expenditure (MET) from the accelerometer data which showed correlation greater than 0.88 with energy expenditure observed with a simultaneous VO2 measurement. Eight healthy subjects wore the accelerometer on waist and simultaneously were breathing through a breathing gas analyzer (Cosmed srl, Italy). The subjects performed eight different one minute tasks with varying intensity including three levels of walking on treadmill. A regression equation was calculated from the measured accelerometer and gas analyzer data. The resulting equation was used to estimate the free living MET from acceleration signals and classify the movement activity information into active ($>$1.5MET) and inactive states ($<$1.5MET). The cut-off point is purposefully fairly low to ensure that the HR data in the inactivity state would not be affected by physical activity. The patients that attended the study were not in good physical condition, so even a light physical activity could increase their HR.

We clustered the data set containing 3 dimensional samples (HR, HRV and movement activity information) to study their characteristics by organizing them into groups based on similarity. Clustering was carried out using k-means algorithm (from *Fuzzy Clustering and Data Analysis Toolbox* for *Matlab*) and more details on the techniques used can be found at *http://www.fmt.vein.hu/softcomp/fclusttoolbox/*. The following measures along with domain knowledge are used to determine optimal number of clusters.

- Partition index (SC): It measures the amount of overlapping between clusters and is defined by the ratio of the sum of compactness and separation of the clusters. This measure is useful when comparing different partitions having equal number of cluster. A lower value of SC indicates a better partition.
- Separation Index (S): It measures the validity of clusters based on minimum distance separation.
- Xie and Beni's Index (XB): It quantifies the ratio of the total variation within clusters and the separation of clusters. The optimal number of clusters will yield minimal value for the index.

III. RESULTS

An example distribution of HR, HRV and MET during one of the recordings from an arbitrarily chosen patient is shown in Figure 1. Initial observations indicate that there is a direct correlation between all the three variables (higher HR, low HRV and high MET). Figure 2 shows the distribution of entire data set over six weeks of rehabilitation period for this participant based on movement activity information (i.e., active and inactive states). This data was from 8 different days during six weeks and it was recorded at least once a week. We can see in the Figure that the means during different days predominantly cluster in 2 zones based on

movement activity information. The ones in active state have high HR and low HRV compared to the inactive state. Next, we will identify the optimal number of clusters and their implications by using the entire data set from all participants.

A. Optimal number of clusters

The data set had 18202 samples for HR, HRV and movement activity information (viz., around 300 hours spread over 6 weeks and 7 patients). K-means clustering algorithm was applied on this data several times with different initial values. There was no clear indication of optimal number of clusters based on local minima for SC, S and XB. It varied from 4 to 6 in most of simulations. Figure 3 shows the mean values for SC, S and XB based on 100 simulations. SC and S almost reach minima at c = 5 point. The XB index reaches this local minimum at c = 4. The optimal number of clusters could be either 4 or 5. We chose 4 as the optimal number since all 4 of them had physiological significance in cardiac rehabilitation domain.

Fig. 3. Values of Partition Coefficient (PC), Separation Index (S) and Xie and Beni's Index (XB) for number of clusters = 3 to 8. The optimal number of clusters is 4 based on minima.
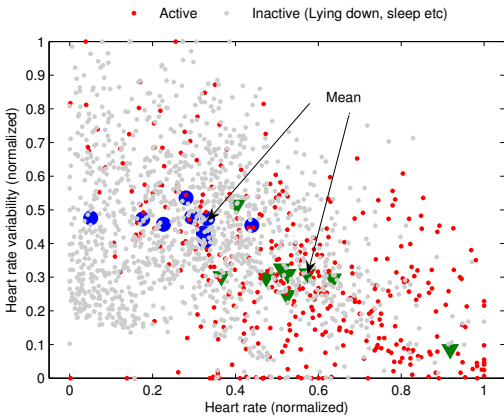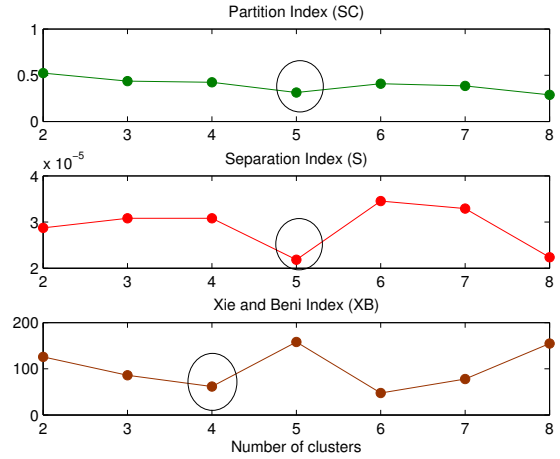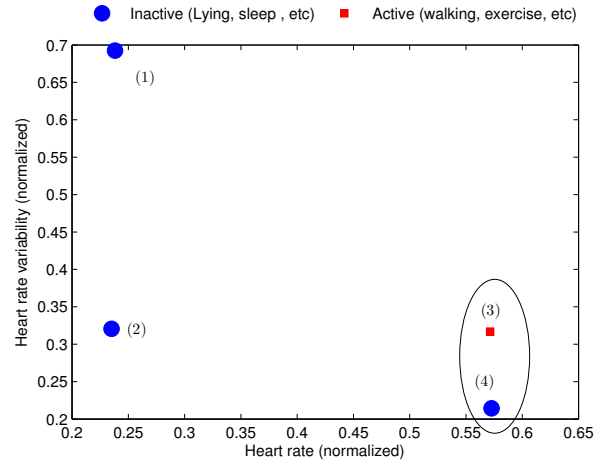

Fig. 2. Distribution of HRV and HRV for an arbitrarily chosen participant (grey = inactive state and red = active state). The means for different days in the data for active and inactive states are shown in blue circles and green triangles respectively.

Fig. 4. Centroid for 4 different clusters.

B. Cluster significance

The centroids for 4 different clusters are shown in Figure 4 and the data distribution based on the clusters is presented in Figure 5 (Statistics for each patient along with their medical conditions are tabulated in Table I). There is only one cluster in active state and the rest are in inactive state. Each of these clusters have characteristic features with physiological and health implications. Some key features of clusters (1), (2), (3) and (4) are -.

- Cluster (1) which falls in inactive state (<1.4 MET) has low HR and high HRV (Centroid = [0.2382 0.6925]). This is a characteristic feature of relaxed state and good sleep induced by increased parasympathetic control.
- Both Cluster (3) and cluster (4) have similar characteristic features of high HR and low HRV. Increased HR and/or reduced HRV is triggered by sympathetic activity. As discussed earlier, sympathetic activity could
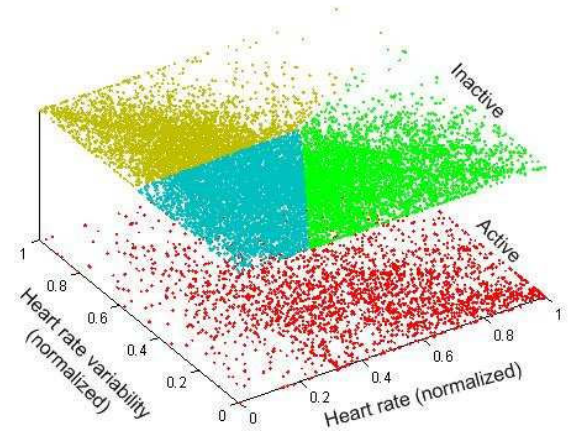

Fig. 5. Data points color coded based on the cluster to which they belong.

TABLE I

AGE, BODY MASS INDEX (BMI), MEDIAN DISTANCE FROM CLUSTERS (1) -(4) WITH (STANDARD DEVIATION), DIAGNOSIS OF DEPRESSION, CARDIAC DISEASE AND β-BLOCKER USAGE FOR EACH SUBJECT. [N/A = NOT AVAILABLE]

Age	BMI	(1)	(2)	(3)	(4)	Depression	Condition	β-blocker
62	33.14	1.03(0.25)	1.02(0.27)	0.44(0.32)	1.05(0.35)	$\sqrt{}$	Valve block	$\sqrt{}$
59	28.96	1.04(0.19)	1.03(0.21)	0.42(0.26)	1.08(0.25)	$\sqrt{}$	coronary artery disease	$\sqrt{}$
70	24.73	1.02(0.21)	1.05(0.24)	0.47(0.29)	1.07(0.33)	X	myocardial infarction	n/a
61	25.36	1.08(0.21)	1.02(0.22)	0.38(0.26)	1.05(0.26)	n/a	coronary artery disease	$\sqrt{}$
58	34.28	1.03(0.23)	1.02(0.25)	0.42(0.29)	1.07(0.30)	X	stent metalyse	$\sqrt{}$
46	18.85	1.07(0.22)	1.02(0.25)	0.29(0.29)	1.03(0.28)	n/a	myocardial infarction	$\sqrt{}$
73	n/a	1.12(0.25)	1.04(0.25)	0.35(0.30)	1.03(0.29)	n/a	n/a	n/a

be induced by physical activity or due to stress, fatigue, depression or other psychological reasons. The sympathetic activity in cluster (3) is activated by physical activity since it is in active state (>1.4MET). Cluster (4) with characteristics of high sympathetic activity in inactive state could indicate the presence of stress and fatigue. Movement activity information is essential to differentiate between these two clusters.

- Cluster (2) which again falls in inactive zone (<1.4 MET) has both low HR and low HRV. Low HRV in inactive state indicates increased sympathetic activity. This cluster needs further clinical investigations and analysis.

IV. DISCUSSION

We identified 4 characteristic clusters based on HR, HRV and movement state information in cardiac patients. 3 clusters are observed in inactive state and 1 in active state. The cluster in active state is characterized by high HR and low HRV due to increased sympathetic activity caused by exercise. The presence of a cluster with high HR and low HRV in inactive state indicates that the sympathetic activity could be triggered by stress and fatigue. Movement activity information plays a critical role in differentiating these clusters. Increased parasympathetic activity leading to relaxed state in inactive state was characterized by cluster with high HRV and low HR. We also observed a cluster in inactive state with both HR and HRV reduced. Reduced HRV in this cluster indicates higher sympathetic activity but the presence of low HR at the same time needs further investigation. These results suggest that by clustering HR, HRV and movement activity data it is possible to discriminate between normal and abnormal physiological state. A difference in distribution between the obtained clusters could indicate an abnormal autonomic function.

In this work, we have broadly classified movement activity information into inactive and active states based on MET thresholds. It is worth noting that discrimination based on 2 state movement activity information may not be adequate since there is a large scatter of data in active zone. The clustering quality can be further improved by including additional informative features. Some of the features are actual MET levels, behavioral information, perceived effort level (Borg Scale), stress scores, medication history, risk factors, circadian patterns, etc. This clustering was done using data

from cardiac patients. It would be good to have data from healthy subjects for comparison and disease characterization. Also, it would be worthwhile tracking the cluster progression with time in patients undergoing rehabilitation.

REFERENCES

[1] "Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology.", *Circulation*, vol. 93, no. 5, pp. 1043–1065, Mar 1996.

[2] Marjolein H Kamphuis, Mirjam I Geerlings, Jacqueline M Dekker, Simona Giampaoli, Aulikki Nissinen, Diederick E Grobbee, and Daan Kromhout, "Autonomic dysfunction: a link between depression and cardiovascular mortality? the fine study.", *Eur J Cardiovasc Prev Rehabil*, vol. 14, no. 6, pp. 796–802, Dec 2007.

[3] P. J. Schwartz, E. Vanoli, M. Stramba-Badiale, G. M. De Ferrari, G. E. Billman, and R. D. Foreman, "Autonomic mechanisms and sudden death. new insights from analysis of baroreceptor reflexes in conscious dogs with and without a myocardial infarction.", *Circulation*, vol. 78, no. 4, pp. 969–979, Oct 1988.

[4] P. J. Schwartz, M. T. La Rovere, and E. Vanoli, "Autonomic nervous system and sudden cardiac death. experimental basis and clinical observations for post-myocardial infarction risk stratification.", *Circulation*, vol. 85, no. 1 Suppl, pp. I77–I91, Jan 1992.

[5] J. T. Bigger, J. L. Fleiss, L. M. Rolnitzky, and R. C. Steinman, "The ability of several short-term measures of rr variability to predict mortality after myocardial infarction.", *Circulation*, vol. 88, no. 3, pp. 927–934, Sep 1993.

[6] Tuomo Nieminen, Mika Khnen, Tiit Kbi, Kjell Nikus, and Jari Viik, "Heart rate variability is dependent on the level of heart rate.", *Am Heart J*, vol. 154, no. 1, pp. e13; author reply e15, Jul 2007.

[7] Klaus Schmid, Yvonne Riehm, Bernd Rossbach, Stephan Letzel, Hans Drexler, and Michael Mck-Weymann, "[influence of rescue service personnel's shift work on psycho-physical parameters.]", *Psychother Psychosom Med Psychol*, Dec 2007.

[8] Vegard Bruun Wyller, Riccardo Barbieri, Erik Thaulow, and J. Philip Saul, "Enhanced vagal withdrawal during mild orthostatic stress in adolescents with chronic fatigue.", *Ann Noninvasive Electrocardiol*, vol. 13, no. 1, pp. 67–73, Jan 2008.

[9] Roumiana S Boneva, Michael J Decker, Elizabeth M Maloney, Jin-Mann Lin, James F Jones, Helgi G Helgason, Christine M Heim, David B Rye, and William C Reeves, "Higher heart rate and reduced heart rate variability persist during sleep in chronic fatigue syndrome: a population-based study.", *Auton Neurosci*, vol. 137, no. 1-2, pp. 94–101, Dec 2007.

[10] Suzanne C Segerstrom and Lise Solberg Nes, "Heart rate variability reflects self-regulatory strength, effort, and fatigue.", *Psychol Sci*, vol. 18, no. 3, pp. 275–281, Mar 2007.

[11] Kaviraja Udupa, T. N. Sathyaprabha, Jagadisha Thirthalli, K. R. Kishore, G. S. Lavekar, T. R. Raju, and B. N. Gangadhar, "Alteration of cardiac autonomic functions in patients with major depression: a study using heart rate variability measures.", *J Affect Disord*, vol. 100, no. 1-3, pp. 137–141, Jun 2007.

[12] Gian Domenico Pinna, Roberto Maestri, Antoni Torunski, Ludmila Danilowicz-Szymanowicz, Malgorzata Szwoch, Maria Teresa La Rovere, and Grzegorz Raczak, "Heart rate variability measures: a fresh look at reliability.", *Clin Sci (Lond)*, vol. 113, no. 3, pp. 131–140, Aug 2007.