Phenotyping Cardiovascular Patients via PPG Signal Clustering Using Symmetric Projection Attractor Reconstruction (SPAR)

Mikołaj Basza[1], Mateusz Soliński[2], Weronika Kowalczyk[3], Damian Walag[4], Łukasz Kołtowski[3]

[1] Medical University of Silesia in Katowice, Poland
[2] Department of Engineering and School of Biomedical Engineering & Imaging Sciences, King's College London, London, UK
[3] Medical University of Warsaw, Warsaw, Poland
[4] Center For Digital Medicine, National Institute of Cardiology, Warsaw, Poland

Abstract

Photoplethysmography (PPG) is increasingly applied in cardiology beyond heart rate and SpO_2, including atrial fibrillation (AF) classification, yet links between other cardiac parameters and PPG morphology remain unclear. In this study, we used the Symmetric Projection Attractor Reconstruction (SPAR) on 10-minute, resting-state PPG signals obtained from 94 cardiovascular patients (mean age 68.9±12.6, 47 females) by the MAX30102 fingertip sensor. Routine clinical data, including echocardiography, were collected. SPAR attractor density profiles were compared using Euclidean distance, and hierarchical clustering was used to group patients. In the entire cohort, we found three clusters, and AF %prevalence was the most significant factor that differentiated them (p=0.018). In the subgroup without AF (n=68), four clusters were found and 5 clinical variables showed significant differences: EF<50% (p=0.034), Age (p=0.035), diabetes (DM2; p=0.039), and Aortic Stenosis (AS; p=0.043). Cluster 1 appeared the 'healthiest', with low prevalence of reduced EF<50%, DM2, and AS. In contrast, Cluster 2 showed the highest proportion of patients with EF<50%, Cluster 3 was enriched in DM2, and Cluster 4 included older patients with frequent AO. Our results shows the potential of using SPAR on PPG in screening cardiac dysfunctions using wearable devices, including heart failure.

1. Introduction

Photoplethysmography (PPG) is finding new applications in cardiology far beyond heart rate and SpO_2 measurements, especially in the classification of atrial fibrillation (AF) and other heart rhythm irregularities. In addition to analysing the inter-beat intervals (IBI), there has been growing interest in analysing PPG waveform morphology, as subtle changes in its shape can provide insights into underlying hemodynamic processes and cardiac function. Morphological analysis enables the detection of beat-to-beat variations associated with vascular compliance, pulse wave dynamics, and contractility [1]. These features may be particularly informative in detecting and characterising cardiovascular dysfunction. One promising approach is the Symmetric Projection Attractor Reconstruction (SPAR) method, which visualises the trajectories of successive beats, capturing both recurring morphological patterns and their variability [2]. This method was used previously in analysing PPG signals or blood pressure (BP) waveforms in the extraction of respiratory rate [3] or systolic peak detection [4]. The main goal of our study is to utilise SPAR in phenotyping cardiovascular patients, varied in hemodynamic and structural heart parameters. The motivation for this analysis is to explore the associations between PPG and echocardiography data, adjusted by demographic information that would extend the possible application of the PPG technology, particularly in wearables.

2. Methods

2.1. Data collection

In this observational study, we obtained data from a cohort of patients admitted to a cardiology ward between 2021 and 2023 and recruited for 10-minute PPG signal measurement. The signals were collected at rest under standardised conditions using red light with the MAX30102 fingertip sensor. The PPG data were matched with routine clinical data, including demographics and echocardiography obtained from the Electronic Health Records (EHR). The list of the parameters and indicators includes: age, sex, BMI, BSA (Body Surface Area), DM2 – type 2 diabetes, EF – ejection fraction, HFpEF, HFmrEF, HFrEF - Heart failure with preserved (EF>60%), mildly reduced (EF: 50-60%), reduced (EF<50%) ejection

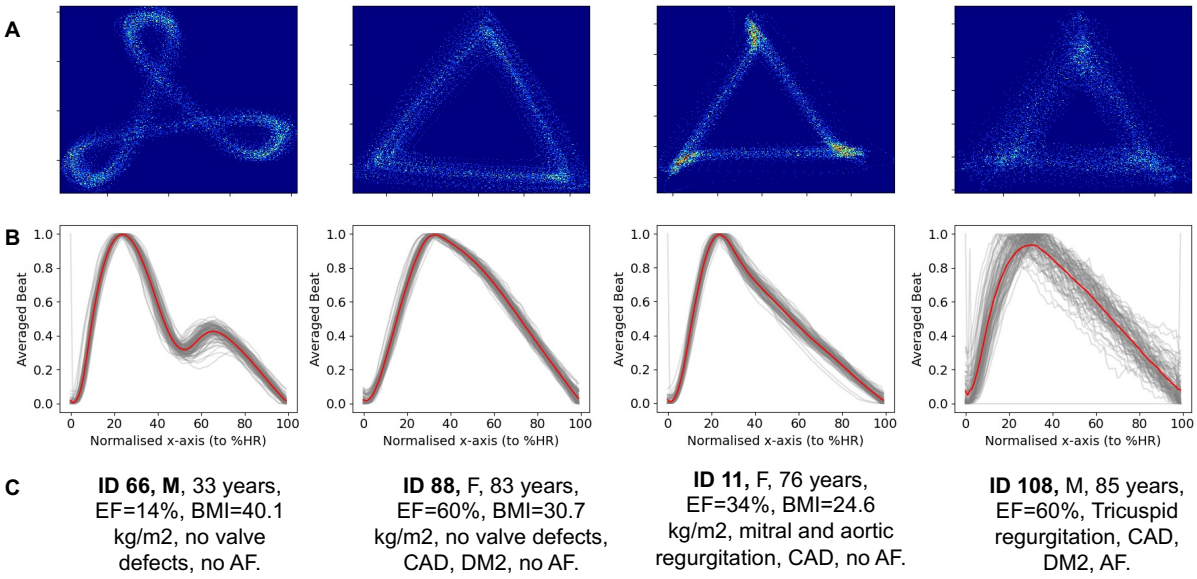

Figure 1. A. Examples of attractors obtained using the SPAR method on PPG waveforms. B. All beats centred to their onset and the averaged beat (red). C. Patients' clinical information.

fraction, aortic/mitral stenosis, aortic/mitral regurgitation, CAD – coronary artery disease, smoking status, TAPSE - Tricuspid annular plane systolic excursion (in mm), AcT - Activated clotting time (in s), LVDD - Left ventricular diastolic diameter (in mm), RV - right ventricular size (mm), Ao - aorta diameter (mm), PWDT - Posterior Wall diastolic Thickness (mm).

2.2. PPG Data processing

The PPG signals were pre-processed before the attractor reconstruction procedures under the following steps:

A) Signal quality evaluation: 10-minute signals were evaluated using a modified signal quality algorithm based on [5]. We modified some of the rules described in the original work and extended the moving window length to 5 s with a hop size of 1 s. Each window was sequentially labelled as bad/good quality. Additionally, we did not take to analysis windows considered as a good quality but surrounded by bad quality periods. **B)** Fiducial points detection: We detected systolic peaks (sp) and beats onsets (on) using the MSPTDfast PPG beat detection algorithm [6]. Using these indices, we calculated a perfusion index (PI) in each beat. **C)** Selection of the 90-second fragment for analysis: from each 10-minute recording, we calculated the mean PI [mean(PI)] and its standard deviation [std(PI)] and the percentage of good quality classifications in 90-second moving windows (with a hop size of 1s). We selected one window with at least 80% of the good quality fragments and has a minimum root square difference of the mean(PI) to the maximum mean(PI) and the std(PI) to the minimum std(PI). These criteria ensure that only the fragments with the highest quality, maximum perfusion and maximum stability were taken for analysis from each patient. **D)** Trend removal: We removed trends from the raw PPG signals using the moving mean with a window of 100 samples. **E)** Time normalisation: we normalised the length of each beat to 100 samples and created a new time series. This normalisation standardises the attractor obtained from signals with different heart rates.

2.3. SPAR

We used the SPAR method to reconstruct attractors from the time-normalised windows of PPG signals. The full description of the method is to be found here [2]. Briefly, the SPAR is realised by the following steps:

1. Create N=3 variables, x, y, and z, that take values from the PPG signal from the moving points separated by a time delay (1/N of the average inter-beat interval).

2. An iteration of the xi, yi, zi creates a single point in an N-dimensional phase space. A collection of all data points produces a trajectory of the signal dynamics.

3. All data points of the attractor are rotated to the standard position.

4. The attractor is mapped into 2D plane.

5. A density "heat map" is created to highlight areas where individual loops overlap.

The heatmaps are used to create density profiles in polar coordinates:

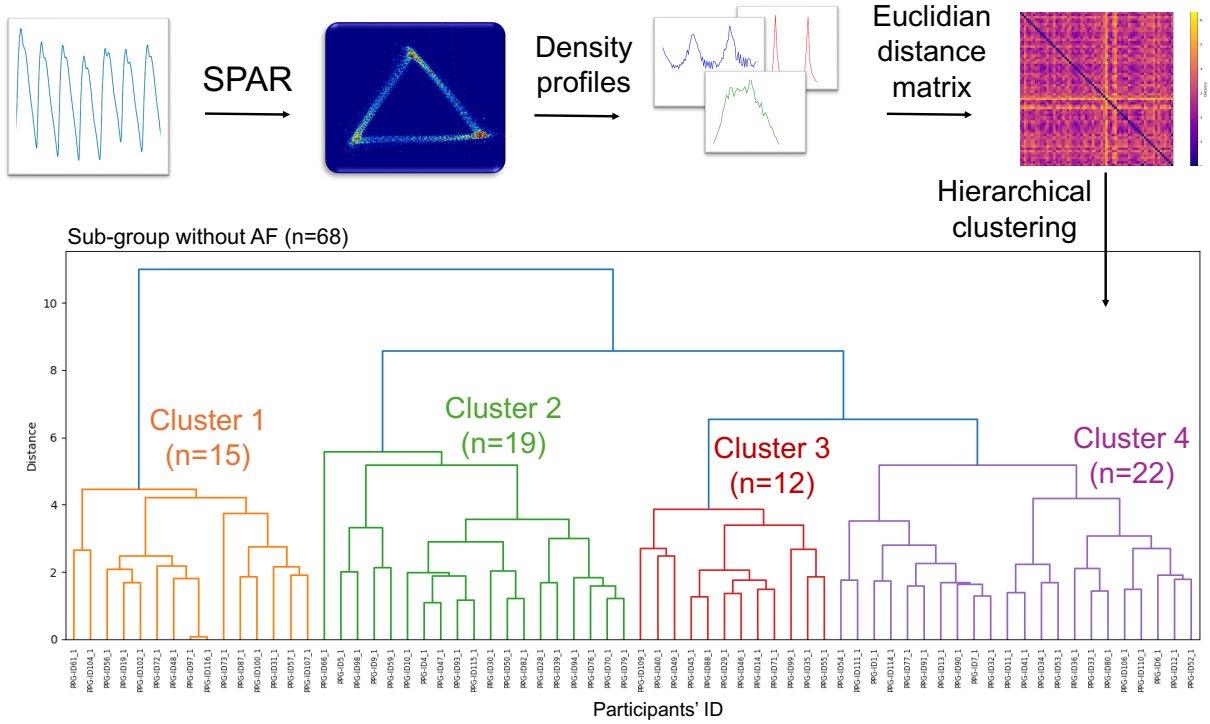

Figure 2. The pipeline of performing hierarchical clustering on PPG signals followed by a dendrogram depicting the clustering results (the subgroup without AF).

• r density distribution (d_r): the density of the attractor as a function of the radius from its centre.

• density distribution (d_θ): the density of the attractor as a function of the angle.

• attractor outline r (d_o): the maximum r of the attractor for a given angle.

2.4. Clustering & phenotyping

Clustering was performed using merged density profiles [d_r, d_θ, d_o] from each attractor. Euclidean distances between profiles served as input to hierarchical clustering with Ward's linkage. Clinical and hemodynamic variables were compared across clusters using chi-squared tests (with Yates' correction) for binary data and ANOVA or Kruskal–Wallis tests for continuous data.

3. Results

We obtained PPG signals from 108 patients, among whom 94 had good-quality signals for further analysis. The mean age of that cohort was 67.5±12.7, consisting of 43 females (46%) and 26 patients (28%) with AF. It is known that this arrhythmia can significantly change the dynamics of pulse wave morphology [7]. Therefore, we performed clustering either on the entire cohort or in a subgroup without AF (n=68). The examples of the attractors extracted from the PPG signals are presented in Figure 1.

The results on clustering of the entire cohort (n=94) revealed three clusters, with AF prevalence as the most significant factor (p = 0.018), followed by HFmEF (p = 0.020) and RV (p = 0.047).

In the subgroup of patients without AF (n=68), four distinct PPG morphology clusters were identified (Figure 2) and 5 significant variables that differentiate them: HFmEF, EF<50%, Age, DM2, and Aortic Stenosis (AS). The comparison of the variables with significant differences between clusters created either for the entire cohort or a subgroup is shown in Table 1.

4. Discussion

SPAR reveals substantial heterogeneity in PPG waveform patterns across patients. A major driver of attractor shape is the dicrotic notch: a deep notch, common in younger individuals, produces looping around the three attractor corners. That number of corners is due to using N=3-dimensional space, which was previously used for PPG [2]. A future direction is to select the embedding dimension by maximising mutual entropy (according to Takens' theory), which may yield a more faithful view of PPG complexity. The differences in attractor properties are also observed considering the prevalence of AF. The beat-to-

Table 1. List of the significant clinical parameters between clusters based on PPG signal morphology.

A. Entire cohort (n=94)				
Variable	**Cl 1 (n=30)**	**Cl 2 (n=29)**	**Cl 3 (n=35)**	p-value
AF [%prev.]	43.3%	10.3%	28.6%	0.018
HFmEF [%]	10.0%	20.7%	0.0%	0.020
RV [cm]	3.22	2.93	2.99	0.047

B. Patients without AF (n=68)					
Variable	**Cl 1 (n=15)**	**Cl 2 (n=19)**	**Cl 3 (n=12)**	**Cl 4 (n=22)**	**p-value**
HFmEF (%)	13.3%	26.3%	0%	0%	0.025
EF<50%	13.3%	57.9%	25.0%	27.3%	0.034
Age [years]	65.2	61.8	58.3	70.8	0.035
DM2 [%]	6.7%	10.5%	41.7%	36.4%	0.039
AS [%]	6.7%	0%	0%	22.7%	0.043

beat fluctuation due to the irregularity of the rhythm in patients with AF introduces a high level of dispersion of the data points in the attractor.

Significant differences were observed in variables reflecting cardiac structure and function: HFmEF (entire cohort and subgroup), RV, age, and EF<50% (subgroup). Although cluster prevalences often fell below 50%, distinct phenotypes emerged in the AF-free subgroup ($n = 68$): Cluster 1 appeared "healthiest" (low reduced EF, DM2, and AS); Cluster 2 had more reduced EF; Cluster 3 was enriched for DM2; and Cluster 4 comprised older patients with frequent AS. Despite these differences, visual separation of attractors is challenging due to high inter-individual variability in density profiles. A practical remedy is to compute partial differences on merged density profiles and localise the most discriminative fragments.

Differences in HFmEF and EF$< 50\%$ support PPG's potential for heart-failure screening. Prior work distinguished healthy vs HF using PPG-derived HRV [8]; our findings indicate that morphology dynamics themselves may be diagnostically useful. Further studies should map specific morphologies linked to HF.

5. Conclusions

We presented the results of an explanatory study about the phenotyping of cardiovascular patients using the PPG signal and transformed into attractors by the SPAR method. It shows the potential of PPG in triage and screening cardiac dysfunctions using wearable devices, including heart failure. With larger datasets and clinical validation, the SPAR method could provide a valuable input to phenotyping algorithms and serve as an alternative to feature-based approaches.

Acknowledgments

MS is part of the COSMOS and HEART.FM projects, which have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement Nos. 788960 and 957532).

References

[1] Charlton PH, Paliakaitė B, et al. Assessing hemodynamics from the photoplethysmogram to gain insights into vascular age: a review from vascagenet. American Journal of Physiology Heart and Circulatory Physiology April 2022; 322(4):H493–H522. ISSN 1522-1539.

[2] Nandi M, Aston PJ. Extracting new information from old waveforms: Symmetric projection attractor reconstruction: Where maths meets medicine. Experimental Physiology May 2020;105(9):1444–1451. ISSN 1469-445X.

[3] Rameshwaran AK, Jayaraman S. Phase space reconstruction for noise-robust respiration rate estimation from ppg signal. In 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, July 2024; 1–4.

[4] Pettit C, Charlton PH, Aston PJ. Photoplethysmogram beat detection using symmetric projection attractor reconstruction. Frontiers in Physiology February 2024;15. ISSN 1664-042X.

[5] Vadrevu S, Manikandan MS. Real-time ppg signal quality assessment system for improving battery life and false alarms. IEEE Transactions on Circuits and Systems II Express Briefs November 2019;66(11):1910–1914. ISSN 1558-3791.

[6] Charlton PH, Argüello-Prada EJ, Mant J, Kyriacou PA. The msptdfast photoplethysmography beat detection algorithm: design, benchmarking, and open-source distribution. Physiological Measurement March 2025;46(3):035002. ISSN 1361-6579.

[7] Basza M, Walag D, Kowalczyk W, Bożym A, Ciurla M, Krzyżanowska M, Maciejewski C, Bojanowicz W, Soliński M, Kołtowski Photoplethysmography wave morphology in patients with atrial fibrillation. Physiological Measurement April 2023;44(4):045001. ISSN 1361-6579.

[8] Al Younis SM, Hadjileontiadis LJ, Stefanini C, Khandoker AH. Non-invasive technologies for heart failure, systolic and diastolic dysfunction modeling: a scoping review. Frontiers in Bioengineering and Biotechnology October 2023;11. ISSN 2296-4185.

Address for correspondence:

Mikołaj Basza,
Medical University of Silesia in Katowice, Poland
E-mail address: baszamikolaj@gmail.com