Spatial Correlation Communities in Panoramic Optical Mapping

Estela Sánchez-Carballo[1], Francisco M Melgarejo-Meseguer[1], Jimena Siles-Paredes[2], Saleem-Ullah[2], Angélica Quadros[2], João Salinet[2], José Luis Rojo-Álvarez[1]

[1] Department of Signal Theory and Communications, Telematics and Computing Systems, Universidad Rey Juan Carlos, Madrid, Spain
[2] HEartLab, Federal University of ABC, São Bernardo do Campo, Brazil

Abstract

Panoramic optical mapping provides rich spatiotemporal information on cardiac electrical activity through video recordings, where each pixel corresponds to a signal evolving over time. This complete information can help characterize different electrical behaviors within the heart tissue, however, analyzing such a large amount of data efficiently is challenging. In this study, we propose a novel methodology for analyzing all the signals simultaneously and also considering their spatial location. Our approach reduces the temporal dimensionality of the signals using manifold learning and then identifies groups of signals using community detection (Louvain algorithm) and clustering (K-means). These two techniques were compared both qualitatively and quantitatively using the modularity index. For Subject 1 (sinus rhythm), Louvain identified 16 communities with a modularity of 0.8763, while K-means identified 14 clusters with a modularity of 0.8492. For Subject 2 (atrial tachycardia), Louvain found 12 communities (modularity 0.8282) and 15 K-means clusters (modularity 0.8049). Louvain always achieved higher modularity, suggesting more coherent community structures. This methodology enables visualization of thousands of signals in a latent space, and may reveal organizational patterns in the heart, highlight pre-processing requirements, and provide insight into the propagation of electrical impulses.

1. Introduction

Cardiac arrhythmias represent a significant health problem in our society, affecting more than 60 million people worldwide. Depending on the type of arrhythmia, patients can develop serious complications such as stroke, heart failure, or even sudden cardiac death if the condition is not recognized and treated immediately [1]. Therefore, it is fundamental to study and develop new approaches that enable a complete understanding of arrhythmias, including their underlying mechanisms and their spatiotemporal organization within the heart.

Panoramic optical mapping (POM) involves acquiring images of the heart from multiple viewpoints using multiple optical cameras, where each camera extracts information from a specific heart region. Therefore, this imaging technique provides fluorescent recordings of the electrophysiological activity on the entire epicardial surface, providing detailed information on the spatiotemporal propagation of electrical activity responsible for maintaining cardiac arrhythmias [2].

POM provides spatio-temporally rich information in the form of a video, where each pixel corresponds to a signal that evolves over time. Therefore, a complete dataset for a single subject can consist of thousands of signals. Manifold learning (MnL) techniques can help analyze such large datasets by reducing their temporal dimensions to two or three, enabling the visualization of all information simultaneously in a latent space that preserves the most relevant features of the original data. Furthermore, the embedded space allows for further analyses, such as identifying groups of signals with shared characteristics, which can reveal regions of abnormal electrical conduction and other electrophysiological phenomena [3].

In this study, we propose a methodology that allows a comprehensive analysis of the information provided by POM, combining MnL with a comparison of two different group detection techniques that can regionalize heart areas and allow the study of their behavior. The structure of this work is as follows. The methodology is detailed in Section 2, the experiments performed and the results obtained are presented in Section 3, and the conclusions are summarized in Section 4.

2. Methodology

The methodology proposed in this work consists of reducing the temporal dimensionality of the data to three using Uniform Manifold Approximation and Projection (UMAP) [4], identifying communities in the embedded space with the Louvain algorithm [5] or clusters with K-

ISSN: 2325-887X DOI: 10.22489/CinC.2025.162

means [6], and reconstructing the original image. The three methods are summarized in this Section.

UMAP is a MnL method that projects high-dimensional data into a latent space while preserving both local and global structures. It achieves this by constructing a high-dimensional graph representation of the data and then optimizing a corresponding low-dimensional graph to be as structurally similar as possible to its high-dimensional counterpart. In this framework, UMAP characterizes the similarity between two points $\mathbf{x_i}$ and $\mathbf{x_j}$ in the original space as probabilities (p_{ij}), and also between their corresponding embeddings $\mathbf{y_i}$ and $\mathbf{y_j}$ in the lower dimensional space (q_{ij}). Then, the binary cross entropy (BCE) between p_{ij} and q_{ij}, i.e.,

$$C = \sum_i \sum_j p_{ij} \log\left(\frac{p_{ij}}{q_{ij}}\right) + (1-p_{ij}) \log\left(\frac{1-p_{ij}}{1-q_{ij}}\right), \tag{1}$$

is minimized using the stochastic gradient descent algorithm. The first term of the BCE function encourages the embeddings of neighboring points to move closer together, and is activated when $\mathbf{x_i}$ is a neighbor of $\mathbf{x_j}$, or vice versa, or when both points are neighbors. In contrast, the second term repels the embeddings of non-neighboring points, pushing them farther apart.

The Louvain algorithm uses the UMAP-optimized graph to detect nonoverlapping communities by maximizing the graph modularity (GM), defined as follows.

$$GM = \frac{1}{2\cdot m} \cdot \sum_{ij} \left(\left[A_{ij} - \frac{k_i \cdot k_j}{2\cdot m} \right] \cdot \delta(c_i, c_j) \right) \tag{2}$$

where m is the number of edges, A_{ij} indicates whether the nodes i and j are connected, k_i is the degree of the node i, and $\delta(c_i, c_j)$ equals 1 if the nodes i and j belong to the same community and 0 otherwise. The Louvain algorithm consists of two phases. During the local optimization phase, nodes are reassigned to neighboring communities if this increases GM, and during the aggregation phase, nodes in the same community are merged into a single node to form a reduced graph. These phases repeat iteratively until GM can no longer be improved.

K-means is a clustering algorithm that divides a dataset into K non-overlapping groups by minimizing the variance within the cluster. It organizes data points $\{\mathbf{y_1}, \mathbf{y_2}, \ldots, \mathbf{y_n}\}$ into K clusters with centroids $\{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_K\}$. The objective function to minimize is defined as follows,

$$J = \sum_{k=1}^{K} \sum_{\mathbf{x_i}\in\mathbf{C_k}} \|\mathbf{x_i} - \boldsymbol{\mu}_\mathbf{k}\|^\mathbf{2}, \tag{3}$$

where C_k is the set of points assigned to the group k, and $\boldsymbol{\mu}_k$ is the centroid of that group. The algorithm consists of two steps, that is, the assignment step, where each point $\mathbf{x_i}$ is assigned to its nearest centroid, and the update step, where each centroid is recomputed as the mean of the points assigned to it. These steps are repeated until the assignments no longer change or the decrease in J becomes negligible.

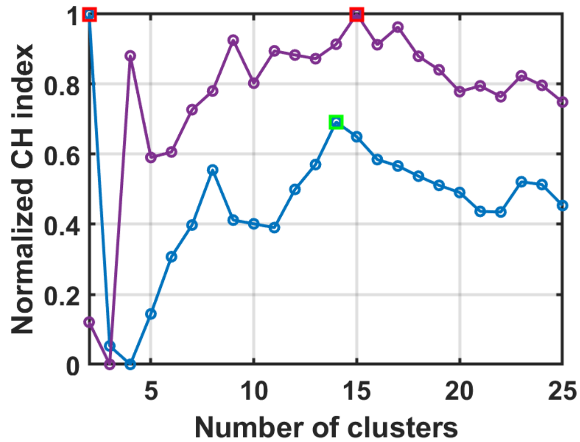

Figure 1: Normalized CH index as a function of the number of clusters for Subjects 1 (blue) and 2 (purple). The global maximum of each curve (red) and the second-highest peak for Subject 1 (green) are marked.

3. Experiments and Results

In this study, we analyzed data obtained from POM experiments performed on isolated rabbit hearts maintained by Langendorff perfusion. The recordings were acquired simultaneously with three cameras, providing multiple views of the epicardial surface. Two subjects were studied: one with sinus rhythm (Subject 1) and one with atrial tachycardia (Subject 2). The complete acquisition protocol is described in [2]. The data were pre-processed in two steps. First, a three-dimensional (3D) median filter with a spatial window of 3×3 pixels and a temporal depth of 5 samples was applied to eliminate high-frequency noise while preserving physiological information. Second, a 4^{th} order high-pass Butterworth filter with a cutoff frequency of 1 Hz was applied in a zero-phase manner to eliminate baseline wander without distorting the signal. After filtering, the data were normalized to the [0,1] range. After preprocessing, UMAP was applied to project each signal into a 3D space. After that, Louvain and K-means were applied to the embedded space and the modularity index was computed to facilitate a quantitative comparison of the grouping results.

One disadvantage of K-means compared to Louvain is that the number of clusters must be specified beforehand. To determine this value, the Calinski–Harabasz (CH) in-

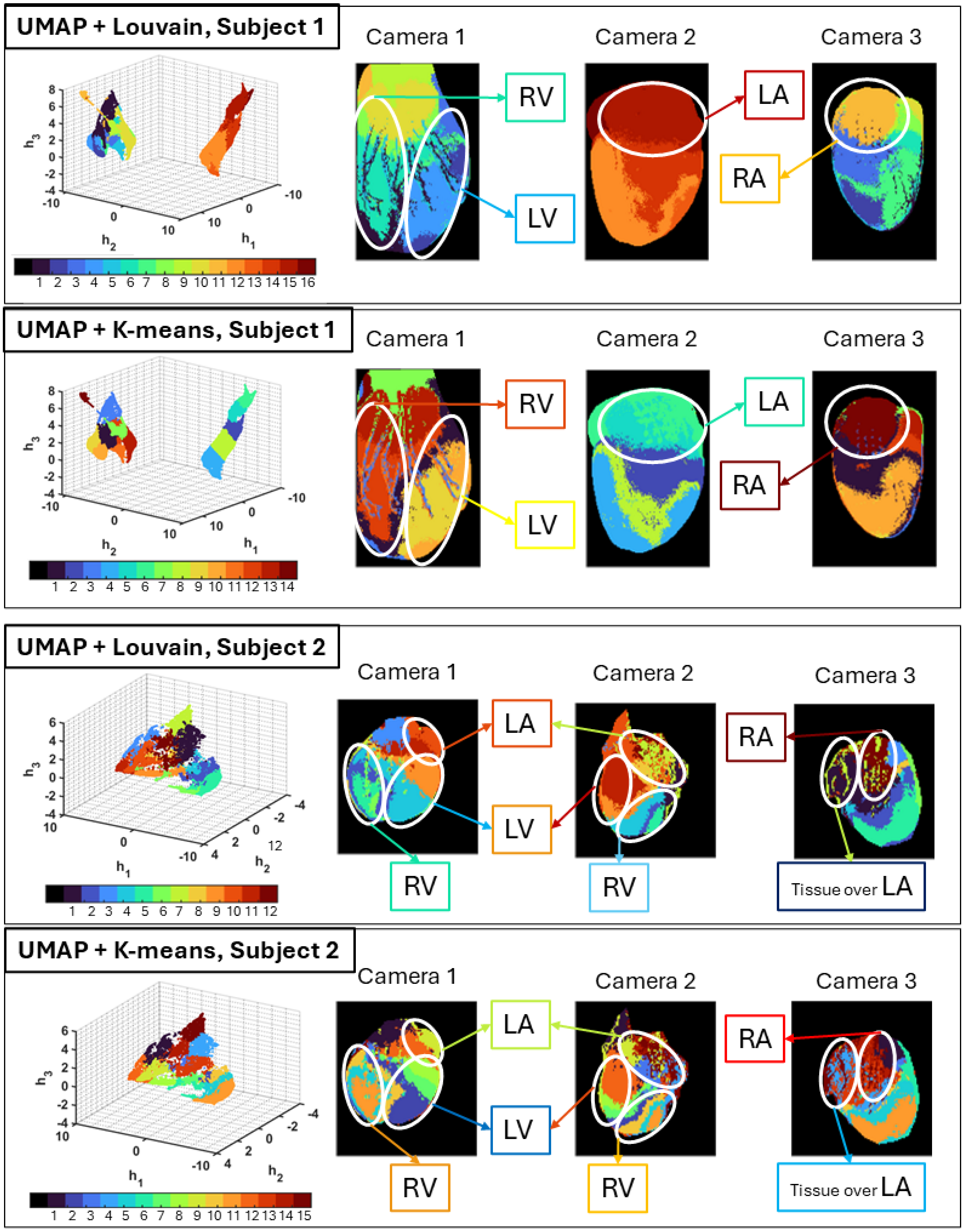

Figure 2: Latent spaces obtained with UMAP (first column), where colors represent communities detected by Louvain (odd rows) or clusters identified by K-means (even rows). Columns two to four show the original image reconstructions for cameras 1, 2, and 3, respectively. Rows 1 and 2 present results for Subject 1, and rows 3 and 4 present results for Subject 2. The colorbar indicates the number of groups identified in each case and their associated colors.

dex was used. Figure 1 shows the normalized CH index for 2 to 25 clusters for Subject 1 (blue) and Subject 2 (purple). The optimal number of clusters for each subject is highlighted in red, corresponding to 2 clusters for Subject 1 and 15 clusters for Subject 2. However, since 2 clusters were insufficient to analyze the structural or mechanistic characteristics of the signals, we selected $K = 14$ for Subject 1, corresponding to the second highest CH index.

Figure 2 shows the complete latent spaces (left) and the original image reconstruction with the colors of the groups encountered (second to last columns). Communities obtained by Louvain are shown in the first and third rows for Subjects 1 and 2, respectively, and clusters obtained by K-means are shown in the second and fourth rows for Subjects 1 and 2, respectively. In Subject 1, there are two well-differentiated clouds of points, one containing the information in cameras 1 and 3, and the other containing the information in camera 2. In this subject, both Louvain and K-means offered similar grouping strategies, with Louvain detecting 16 communities compared to 14 clusters for K-means. The modularity values were 0.8763 for Louvain and 0.8492 for the K-means, indicating that Louvain provided a slightly better community structure. In Subject 2, the ROI was smaller, which may suggest a shared field of view among the different cameras. This overlap may explain why Louvain and K-means identified the signals in the upper part of camera 1, the lower left part of camera 2, and the left part of camera 3 as belonging to the same communities. In this case, Louvain detected 12 communities and 15 clusters of K-means, with modularity values of 0.8282 and 0.8049. In general, both methods generated meaningful groups, with Louvain achieving a higher modularity, suggesting better consistency in the detected grouping structure.

By focusing on the reconstructed images in Fig. 2, it can be observed that some groups lie close to each other, while others are farther apart, suggesting a progression of communities that may reflect the propagation of the electrical impulse across the epicardium. Each group represents a region of the epicardial surface whose signals exhibit a similar behavior, which may be associated with underlying organizational patterns of electrical activity. In addition, some communities appear to contain noisier signals, likely due to a higher level of noise in the original recordings. This, in turn, influences the grouping strategy and can be used to identify which sets of signals require additional preprocessing, as well as the type of preprocessing, to improve subsequent analyses.

4. Conclusions

This work proposes a methodology for analyzing large datasets of high-dimensional spatio-temporal POM signals using MnL and grouping algorithms, i.e., Louvain and K-means. Our approach enables the simultaneous analysis of thousands of recordings from a single patient, providing a spatio-temporal view of epicardial POM signals that may yield insights into organizational patterns in the heart, additional pre-processing requirements, and even the propagation of electrical impulses. When comparing Louvain community detection and K-means clustering, Louvain always achieved higher modularity. It should be noted that partial overlap exists between the cameras fields of view, which may influence the spatial distribution of the detected groups. Future work could integrate electrical recordings that were captured simultaneously to validate and further characterize the groups detected in the optical signals.

Acknowledgments

This work was supported by research Grants PID2022-140786NB-C31, PID2022-140786NB-C32 (LATENTIA), PID2022-140553OA-C42 (PCardioTrials), and PID2023-152331OA-I00 (HERMES) funded by the Spanish Government (/AEI/10.13039/501100011033). Also supported by HERMES 2024/00004/006 (HERMES) from Rey Juan Carlos University, by Madrid Community (ELLIS Madrid Node), and by FAPESP grant #2020/03601-9.

References

[1] Tan SCW, et al. Trends in global burden and socioeconomic profiles of atrial fibrillation and atrial flutter: Insights from the global burden of disease study 2021. CJC Open 2025; 7(3):247–258.

[2] Siles J, et al. An integrated platform for 2-D and 3-D optical and electrical mapping of arrhythmias in langendorff-perfused rabbit hearts. The Journal of Physiology 2025;1–21.

[3] Sánchez-Carballo E, et al. Interpretable manifold learning for T-wave alternans assessment with electrocardiographic imaging. Engineering Applications of Artificial Intelligence 2025;143(109996):1–19.

[4] McInnes L, et al. UMAP: Uniform manifold approximation and projection. Journal of Open Source Software 2018; 3(29):1–2.

[5] Blondel VD, et al. Fast unfolding of communities in large networks. Journal of Statistical Mechanics Theory and Experiment 2008;10(10008):1–12.

[6] Ikotun AM, et al. K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences 2023;622:178–210.

Address for correspondence:

Estela Sánchez–Carballo
Dep. of Signal Theory and Communications
University Rey Juan Carlos, Fuenlabrada (Madrid), Spain
Mail to: estela.sanchezc@urjc.es