Respiratory Physiology & Neurobiology 207 (2015) 28–39

Contents lists available at ScienceDirect

Respiratory Physiology & Neurobiology

journal homepage: www.elsevier.com/locate/resphysiol

Automatic classification of canine PRG neuronal discharge patterns using K-means clustering

Edward J. Zuperku[a,b,*], Ivana Prkic[a,b], Astrid G. Stucke[a,b,c], Justin R. Miller[a,b], Francis A. Hopp[a], Eckehard A. Stuth[a,b,c]

[a] Clement J. Zablocki Department of Veterans Affairs Medical Center, Milwaukee, WI, USA
[b] Department of Anesthesiology, Medical College of Wisconsin, Milwaukee, WI, USA
[c] Children's Hospital of Wisconsin, Pediatric Anesthesia, Milwaukee, WI, USA

ARTICLE INFO

Article history:
Received 20 August 2014
Received in revised form 6 November 2014
Accepted 25 November 2014
Available online 12 December 2014

Keywords:
Classification
Clustering
Discharge patterns
Pontine neurons
Dogs

ABSTRACT

Respiratory-related neurons in the parabrachial-Kölliker-Fuse (PB-KF) region of the pons play a key role in the control of breathing. The neuronal activities of these pontine respiratory group (PRG) neurons exhibit a variety of inspiratory (I), expiratory (E), phase spanning and non-respiratory related (NRM) discharge patterns. Due to the variety of patterns, it can be difficult to classify them into distinct subgroups according to their discharge contours. This report presents a method that automatically classifies neurons according to their discharge patterns and derives an average subgroup contour of each class. It is based on the K-means clustering technique and it is implemented via SigmaPlot User-Defined transform scripts. The discharge patterns of 135 canine PRG neurons were classified into seven distinct subgroups. Additional methods for choosing the optimal number of clusters are described. Analysis of the results suggests that the K-means clustering method offers a robust objective means of both automatically categorizing neuron patterns and establishing the underlying archetypical contours of subtypes based on the discharge patterns of group of neurons.

Published by Elsevier B.V.

1. Introduction

Respiratory-related neurons in the parabrachial-Kölliker-Fuse (PB-KF) region of the pons play a key role in the control of phase timing and breathing frequency (Alheid et al., 2004; Cohen, 1971; Dutschmann and Dick, 2012; Prkic et al., 2012; St-John, 1998). The neuronal activities of these pontine respiratory group (PRG) neurons exhibit a variety of discharge patterns, including inspiratory (I), expiratory (E) and phase spanning patterns (Segers et al., 2008; Song et al., 2006; Ezure and Tanaka, 2006). Also found in this region are non-respiratory related (NRM) neurons, which show tonic activity patterns (Segers et al., 2008). Due to the variety of patterns, it can be difficult to classify them into distinct subgroups according to their discharge contours. The purpose of this report is to present a method that automatically classifies neurons according to their discharge patterns and derives an average subgroup contour of each class. It is based on the K-means clustering technique (Aravind et al., 2010) and it is implemented via SigmaPlot User-Defined transform scripts (SigmaPlot 11.0, Systat Software, Inc. San Jose, CA). In general, clustering involves grouping data into categories based on some measure of inherent similarity or distance.

2. Methods

The discharge patterns of 135 PRG neurons obtained from recordings in a decerebrated dog model ($n = 12$ preparations) were used to develop a method to automatically classify discharge patterns according to their contours. The data were recorded from vagotomized dogs ventilated with an air–O_2 mixture and maintained in hyperoxic isocapnia ($F_{IO2} > 0.6$, end-tidal CO_2 range 40–50 mmHg). Extracellular spike activity was recorded from the PB-KF region using a 16-electrode NeuroNexus probe. The electrodes were linearly arranged with an inter-electrode spacing of 100 μm. The spikes were sorted using Cambridge Electronic Design (CED) Spike2, version 7 software. Timing pulses triggered at the upstroke and post-peak downstroke of the phrenic neurogram were used to create cycle-triggered histograms (CTHs) with 50 or 100 ms bins. The CTHs are expressed in terms of percent of peak discharge frequency (*F*n), since the emphasis is on the contour of the

* Corresponding author at: Research Service/151, Zablocki VA Medical Center, 5000 W. National Avenue, Milwaukee, WI 53295, USA. Tel.: +1 414 384 2000x41579.
E-mail address: ezuperku@mcw.edu (E.J. Zuperku).

http://dx.doi.org/10.1016/j.resp.2014.11.016
1569-9048/Published by Elsevier B.V.

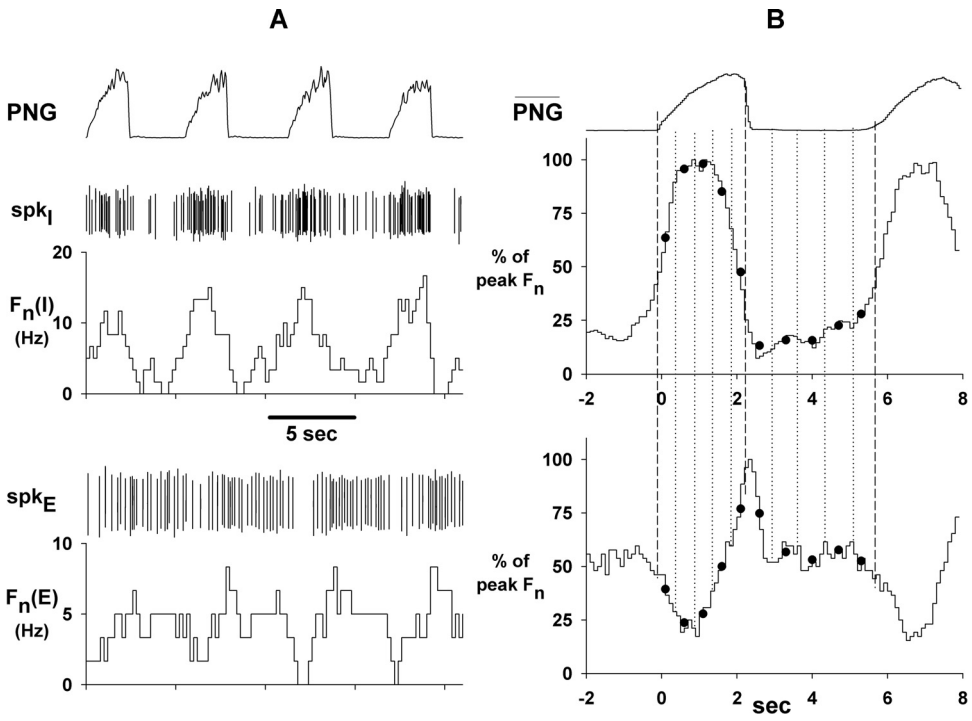

Fig. 1. Two examples illustrating data preparation for the clustering procedure. Panel A: spike activity (CED Spike2 wavemark signals) and corresponding rate meter recordings of an I (upper) and E (lower) PRG neuron. Panel B: Corresponding CTHs and time-averaged phrenic neurogram (PNG-bar; 31 cycles). Dots: 5 temporally equidistant values during I-phase and 5 during the E-phase are selected for each vector representation of neuronal pattern. Time zones are indicated by the vertical dashed and dotted lines. See text for more details.

discharge pattern rather than absolute discharge frequency (amplitude). For each CTH, the I and E phases were each divided into 5 equal zones from which the average discharge frequency (*F*n) was computed, yielding 10 data points (see Fig. 1). Using these data points, each CTH can be represented by a 10-dimensional vector: $\vec{X} = [X_1, X_2, \ldots X_{10}]$. These 135 vectors ($N = 135$ neurons) formed the data set that was then subjected to subgroup assignment (clustering) using a modified K-means method (e.g., Aravind et al., 2010). In preliminary analyses, vector lengths of 8, 12, and 16 were also examined to gain insight into the optimal vector length for best discrimination.

2.1. K-means method

For each cluster (subgroup), an initial pattern (10D vector) is required as the starting point of an iterative procedure. A representative or typical example of a commonly occurring pattern from the data set was selected for each cluster. It is important that each selected pattern is visibly different from the other patterns, otherwise two or more clusters will overlap. In this report seven clusters were designated based on initial patterns such as augmenting and decrementing I and E, IE-EI-phase spanning, and NRM. To determine if an optimal number of clusters was obtained, the data were further examined using a relative distance matrix and a plot (F(k) vs. k) related to the contribution each additional cluster makes to the reduction in overall statistical variance of the data set, adapted from the analysis of Pham et al. (2004) (see Results 3.3).

The next step is to compare each member (neuronal 10-D vector) of the data set, one at a time, to each of the seven selected 10D cluster patterns, $\vec{C}_j = [C_{j1}, C_{j2}, \ldots C_{j10}]$ *for* $j = 1-7$ and determine which cluster, j, is the closest in terms of vector distance. Specifically, the Euclidean distance (D) is calculated (Eq. (1)).

$$D_j = \sqrt{[(X_1 - C_{j1}) + (X_2 - C_{j2}) + \ldots (X_{10} - C_{j10})]^2} \quad (1)$$

where $j = 1$ to 7 is the cluster number and X_i and C_i are corresponding vector components of the data-set member vector and the cluster centroid vector (see below for definition of centroid), respectively. The data vector with the shortest distance to a given cluster is assigned to that cluster and the *shortest* distance is saved and used to calculate the overall iteration error of the complete data set. This procedure is repeated for each member of the data set, in this case $N = 135$. The overall error is the average value of the minimum distances for all N neurons. The next step is to update the C_i values of each cluster. This is done via averaging the data-set member vectors that belong to each cluster (Eq. (2)).

$$\vec{C}_j = \frac{1}{N_j}\sum_{j=1}^{N_j}\vec{X}_j \quad (2)$$

Where N_j is the number of the neuron patterns associated with the j-th cluster. The updated averaged cluster vector is also referred to as the *centroid*. The above procedure is repeated with the updated centroid vectors and the overall minimum error is noted. Typically less than 10 iterations were required for convergence. Appendix A gives the details of the implementation of the procedure using the SigmaPlot application software.

2.2. Weighting of 10 D-vector to reduce effects of outliers on centroid

As a refinement to the clustering method, we have found that applying weighting factors to each of the neuron vectors assigned to a given cluster can reduce the overall error. This process assigns less weight to those vectors that are furthest away from the centroid, such that the shape of the centroid is not distorted by an "outlier" vector. A weighting scheme that works well is a decaying exponential of the form:

$$w_j = \exp[-\alpha(D_j - D_{min})/(D_{max} - D_{min})] \quad (3)$$

Download English Version:

https://daneshyari.com/en/article/2846881

Download Persian Version:

https://daneshyari.com/article/2846881

Daneshyari.com