Abstract

Mouse ultrasonic vocalizations (USVs) are studied in various fields of science. However, background noise and varied USV patterns in observed signals make complete automatic analysis difficult. We propose a series of methods to cluster mouse USV data automatically. The procedure includes noise reduction, detecting USV calls, transforming USV calls as functions and functional clustering. The proposed methods are shown useful with two data sets taken from laboratory mice.

Keywords: breakpoint; B-spline basis function; functional cluster analysis; moving average.

1. Introduction

Recent progress in the studies of ultrasonic vocalization (USV) has revealed a variety of important roles of the USVs in social communication in rodents. Furthermore, important role of different USV patterns have been reported recently. USVs emitted from males have a role to attract females in a play back test. In the study, it has been shown that females prefer to approach towards specific types of USVs emitted from male (Sugimoto et al. 2011), and tend to approach towards USV type that is emitted from different strain from that of her own parents (Asaba et al. 2014). A variety of studies have shown that different patterns of USVs in mice have different effect for social communication. It is important to characterize repertoire of USVs patterns. In order to characterize the pattern of USVs in detail, many steps such as, recording, call detection, call type classification and acoustic analyses, are required by using commercially available software. In these steps, call type classification is still time-consuming, because scientists need to do noise reduction and pattern classification manually in most cases (Scattoni, et al. 2008, Sugimoto, et al. 2011). Currently available software can calculate and express spectrogram and allow automatic signal detection. However, method for automatic classification of USV calls remains to be developed.

In this paper, we propose a series of method to reduce noise, detect USC calls, define USV calls as functions and classify them by a functional clustering method.

2. Noise reduction and detection of USVs

We transform the raw USV voice data into a sequence of USV frames (spectrogram) by FFT. The obtained data is a matrix, denoted by $X = \{x_{i,j}\}$. Here $i = 1, ..., T$ stands for time, $j = 1, ..., F$ indicates frequency. The values $x_{i,j}$ indicate the volume of the vocalization.

The step of noise reduction contains moving average and determination of minimum USV intensity. We use a small matrix with size $m \times n$ to compute the moving average of the original spectrogram matrix.

$$\tilde{x}_{k,l} = \frac{1}{mn} \sum_{i=k}^{k+(m-1)} \sum_{j=l}^{l+(n-1)} x_{i,j}, \quad k = 1, \cdots, T - m + 1, \quad l = 1, \cdots, F - n + 1.$$
This enable us to smooth slight perturbations in the signals. Then we determine the minimum intensity of the USV voice and set all entries less than the intensity to 0. This minimum intensity can be chosen by using the intensity of the background noise.

For each time point, we detect the frequency corresponding to the maximum intensity. The frequency is weighted by the relative intensity $jx_{i,j}/(\sum x_i)$. This reduces the 3-dimentional data into a sequence of 2-dimentional, time and frequency. To detect the USV signals from the sequence, we also determine the length of the interspace between two adjacent USV calls and regard too short USVs as noise by the total number of the USV calls.

3. Definition of functional data
Considering the shape of the USVs, we prefer to use the B-spline basis function method (de Boor, 2000) to define the USVs as functions. For continuous USV calls, we assume that

$$f_{i,j} = \theta_{i,0} + \sum_{p=1}^{P} \theta_{i,p} \beta_p(t_{i,j}) + \varepsilon_{i,j} \tag{1}$$

or in a matrix form

$$f_i = B\theta_i + \varepsilon_i \tag{2}$$

where $t_i = (t_{i,1}, \cdots, t_{i,n})'$, $f_i = (f_{i,1}, \cdots, f_{i,n})'$ stand for time and frequency, respectively. $B = (1, \beta_1(t), \cdots, \beta_P(t))$ is the B-spline basis functions with degree $d$. Using the least square method, for each USV call, we obtain a coefficient vector

$$\hat{\theta}_i = (B' B)^{-1} B' f_i \tag{3}$$

and the regression function

$$\hat{f}_i = B\hat{\theta}_i \tag{4}$$

For discontinuous USV calls, we specify a constant $\kappa$ as threshold. If $|f_{i,j} - f_{i,j-1}| > \kappa$, then the curve is discrete at time $j$. And $j$ is called a breakpoint. To make a jump at a breakpoint in the regression function, we set $d + 1$ knots at the breakpoint. Hence, for each curve, we obtain a coefficient vector with dimension $P = \text{order} + \text{number of interior knots}$.

4. Functional clustering
To make a classification of the obtained USV curves, we first group the curves by the number of breakpoints. That is, curves with same number of breakpoints are gathered into the same cluster, and their coefficient vectors have the same length. These clusters are called the first level clusters.

Secondly, we separate the curves in each first level cluster by their shapes. This process is performed by clustering their coefficient vectors with standard multivariate clustering methods, such as, Ward’s method and K-means method. As shown in Abraham, et al. (2003), this gives the same results as those of functional clustering with the B-spline basis function method.

5. Results
Using the methods in Sections 2 and 3, we obtained 390 USV functions from the data of a BALB/cAnN male mouse. This noise reduction process can be seen from the first 4 panels of Figure 1.

In the 390 USV calls, most of them are continuous. We set $\kappa = 10$ to detect jumps, and obtained 378 continuous USV calls and 12 USV calls with one jump. For continuous USV calls, we equally space 4 interior knots and fit the signals by cubic B-spline functions. For discontinuous USV calls, we set 6 interior knots at each break point and fit the signals by B-spline basis functions with order 5. We do not put other interior knots, because break points occur at different places for different USV calls. To get good fitting, we use higher order splines than cubic splines. The last penal of Figure 1 shows the regression functions of the example signals.
Figure 1: Noise reduction
The final step is the functional clustering. The 390 functions are first classified by their number of break points. Two clusters, continuous curves and discontinuous curves with one break point, are grouped. Because there are only 12 discontinuous curves, we do not classify them further. For continuous curves, we use Ward’s method to obtain the dendrogram as shown in Figure 2. We can divide the 378 continuous curves into 3 or 4 clusters. The k-means method with \(k = 3\) results in the clusters in the first three panels of Figure 3. The last panels is the cluster containing the 12 discontinuous curves. Since we only consider the shape of the curves, we plot all the curves in the interval of time \([0, 1]\).

Figure 2: Cluster dendrogram obtained by clustering the coefficient matrix of continuous functions with Ward’s method.

Figure 3: Clustering of BALB/cAnN data. The first 3 clusters are obtained by 3-means method. The last cluster gathers the discontinuous curves.
References


