Published in final edited form as:

J Neurosci Methods. 2014 October 30; 236: 68–75. doi:10.1016/j.jneumeth.2014.08.007.

Automated Detection of 50-kHz Ultrasonic Vocalizations Using Template Matching in XBAT

David J. Barker^{1,2}, Christopher Herrera¹, and Mark O. West¹

¹Department of Psychology, Rutgers University, 152 Frelinghuysen Road, Piscataway, New Jersey, 08854, USA

²Neural Networks Section, Integrative Neuroscience Research Branch, National Institute on Drug Abuse, 251 Bayview Boulevard, Baltimore, Maryland, USA

Abstract

Background—Ultrasonic vocalizations (USVs) have been utilized to infer animals' affective states in multiple research paradigms including animal models of drug abuse, depression, fear or anxiety disorders, Parkinson's disease, and in studying neural substrates of reward processing. Currently, the analysis of USV data is performed manually, and thus time consuming.

New Method—The goal of the present study was to develop a method for automated USV recognition using a 'template detection' procedure for vocalizations in the 50-kHz range (35-80 kHz). The detector is designed to run within XBAT, a MATLAB graphical user interface and extensible bioacoustics tool developed at Cornell University.

Results—Results show that this method is capable of detecting >90% of emitted USVs and that time spent collecting data by experimenters is greatly reduced.

Comparison with Existing Methods—Currently, no viable and publicly available methods exist for the automated detection of USVs. The present method, in combination with the XBAT environment is ideal for the USV community as it allows others to 1) detect USVs within a user-friendly environment, 2) make improvements to the detector and disseminate and 3) develop new tools for analysis within the MATLAB environment.

Conclusions—The present detector provides an open-source, accurate method for the detection of 50-kHz USVs. Ongoing research will extend the current method for use in the 22-kHz frequency range of ultrasonic vocalizations. Moreover, collaborative efforts among USV

*Corresponding Author: Mark O. West, Ph.D., Department of Psychology, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854, Phone: 848-445-2419, Fax: 732-445-2263, markwest@rutgers.edu.

The authors have no financial interests to be disclosed.

Author Contributions: DJB was responsible for the project design.

DJB, and CH were responsible for data organization and analysis.

DJB and MOW were responsible for writing the manuscript. All authors helped prepare and edit the document.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

researchers might enhance the capabilities of the current detector via changes to the templates and the development of new programs for analysis.

Keywords

Affect; Ultrasonic Vocalizations; Automatic detection; MATLAB; open-source

1.0 Introduction

Rats, like other animals, are capable of intraspecies communication. Specifically, rats produce laryngeal sounds of two distinct types. The first of these types falls within humans' audible range (0-18 kHz) and is produced through slow vibrations of the vocal folds. Sonic vocalizations are elicited by inducing pain or handling naïve subjects in the laboratory (Brudzynski 2008; Litvin et al. 2007). The second type of laryngeal sound that rats produce involves constriction and stabilization of the larynx, such that a small orifice is produced. When air is forced through this orifice, a whistle-like sound—above the human hearing range (Figure 1; 18-80 kHz)—is produced (Johnson et al., 2010). These high-frequency emissions are referred to as ultrasonic vocalizations (USVs; Brudzynski 2008).

Ultrasonic vocalizations have been utilized to infer animals' affective states in multiple research paradigms, including animal models of drug abuse (Panksepp et al. 2002; Covington and Miczek 2003; Thompson et al. 2006; Ahrens et al. 2009; Simola et al. 2010; Barker et al. 2010; Maier et al. 2010; Wright et al. 2010, 2012; Taracha et al. 2012; Meyer et al. 2012), depression, fear or anxiety disorders (Cuomo et al., 1988; Kaltasser, 1991; DeVry et al., 1993; Schreiber et al., 1998; Borta et al., 2006), Parkinson's disease (Ciucci et al., 2007, 2008, 2009), and in studying neural substrates of reward processing (Burgdorf et al., 2001, 2007; Wintink & Brudzynski, 2001).

While USVs can provide an additional and powerful dependent measure to many disease models, the analysis of USV data is currently prohibitive. Specifically, files must be screened manually in order to detect and quantify the observed vocalizations—a process which can take 15-20 hours of human scoring time for every 1 hour of recorded data (Reno et al., 2013). These types of signal-detection problems are common in the bioacoustics field, and viable detectors have been developed for a number of other species (e.g., Brandes, 2008; Mohammad & McHugh, 2011; Chesmore & Ohya, 2004; Charif & Pitzrick, 2008; Mellinger & Clark, 1997). However, to the best of our knowledge, only one method for the automatic detection of USVs is being developed (Reno et al., 2013).

The goal of the present study was to develop a method for automated USV recognition using a 'template detection' procedure for vocalizations in the 50-kHz range (35-80 kHz). The detector is designed to run within XBAT, a MATLAB graphical user interface and extensible bioacoustics tool developed by the Bioacoustics Research Program at Cornell University. Thus, the detector has been customized for use with USVs while retaining a user-friendly interface. In parallel, we have developed additional MATLAB tools (i.e., extensions to XBAT) that are specific to the analysis of USVs and expedite the revision and quantification of detections. Results show that the template library is capable of detecting >90% of emitted USVs and that time spent collecting data by experimenters is greatly

reduced. Moreover, the XBAT environment is ideal for the USV community as its quasi-open source format allows for others in the community not only to use the detector, but also to make adjustments to continually improve the accuracy and generalizability of the detector or to implement new measurement tools for use with USVs within the XBAT environment. Finally, we discuss advantages to using this template detection procedure over other detection methods and explore the realistic limitations of the detector.

2.0 Materials and Methods

2.1 Subjects, Apparatus and Behavioral Procedures

Eighteen USV recordings were taken from male, Long-Evans rats ($n=18$; Charles River, Wilmington, MA) that were catheterized and cared for as described previously (Root et al., 2009, 2011). Subjects were singly housed on a 12h: 12h light dark cycle with dawn at 1030h. Behavioral procedures were conducted in Plexiglas chambers measuring 24 cm $\times$ 34.5 cm $\times$ 34.5cm, which were contained within a larger sound attenuating chamber (~ 76 cm³). Recordings were taken from subjects that had been trained to self-administer cocaine during long-access sessions (6 hours/day) which began daily at 10:30AM. Self-administration training ran 7 days per week for 2-3 weeks. Experimental apparatuses were controlled by a PC running MED-Associates hardware and software (St. Albans, VT). Water was available *ad libitum* except during self-administration sessions. Standard lab chow was provided following self-administration sessions to maintain subjects' weights between 320-340 grams. All protocols were performed in compliance with the Guide for the Care and Use of Laboratory Animals and have been approved by the Institutional Animal Care and Use Committee, Rutgers University.

2.2 USV Recordings

A condenser microphone (CM16/CMPA, Avisoft Bioacoustics, Berlin, Germany) was used for recordings. The microphone was suspended ~ 2.5 cm from a set of small holes in the top of self-administration chambers. Recordings were amplified and digitized by either an Ultrasound Gate 116H (Avisoft Bioacoustics, Berlin Germany), or an Avisoft CM16/CMPA40-5V interface connected to a DT3010/32 analog-to-digital converter (Data Translation, Marlboro, MA). Sonorous activity was recorded at 250 kHz sampling frequency (16-bits) using Avisoft Recorder software (Avisoft Bioacoustics, Berlin Germany) or at a 250-kHz sampling frequency using Sciworks (Datawave Technologies, Loveland, CO). Recorded ".wav" files were analyzed using Raven (Bioacoustics Research Program, 2011; Cornell Lab of Ornithology, Ithaca, NY), which enabled the creation of a spectrograms and the insertion of labels for the frequency and temporal parameters of each call.

2.3 Comprehensive Table Generated by Manual Characterization of USVs

Raven Pro 1.5 (Bioacoustics Research Program, 2011) was used for post-hoc analysis. Each ".wav" file was opened as a spectrogram with a Fast Fourier Transform (FFT) length of 512 samples and a flat-top window with 50% overlap. Spectrograms were visually scanned for patterns resembling USVs. Multiple evaluations of each of the 18 WAV files were conducted by human scorers to create a comprehensive table of all USVs observed within a

given file. This table was designated as the ‘master table’ against which detector results were compared.

2.4 Template Detection

An overview of the steps required to implement template detection is presented in Table 1. Template detection was conducted using signal detection and sound visualization tools in the XBAT environment. USV template libraries (see below) were imported into XBAT's data template detector tool. The data template detector scans recorded sounds and compares them to sets of templates using spectrogram correlation. This allows the detector to quantify the acoustic similarity between sections of the sound file and templates from the target species. Portions of the spectrogram in which the quantified similarity exceeds a specified threshold are stored for subsequent review and analysis as event logs in XBAT. For clarity, USVs scored using template detection will be referred to as ‘detections’ or ‘detected USVs’ throughout, while those scored manually by human observers will be referred to as ‘observations’ or ‘observed USVs’.

2.4.1 Spectrogram Correlation—Spectrogram correlation (Figure 1) works using a sliding window to compare two spectrograms across time. The relationship between the template spectrogram ($X_{t,f}$; t =time; f =frequency) and detection spectrogram ($Y_{t,f}$) produces a correlation value between the two at each different lag (τ). The values being correlated consist of the amplitude values (i.e. spectrogram power) at each frequency bin ($X_{t,f}$) in the fast fourier transform (FFT) at a given lag ($Y_{t+\tau,f}$). These values can be calculated for the entirety of the detection spectrogram by incrementing the value of τ in steps equal to the time grid resolution of the spectrogram. Correlations can range between -1 and 1, with a correlation of 0 representing spectrograms that are orthogonal and a correlation of 1 representing spectrograms that are identical at a given lag τ . The relationship between the template spectrogram and the detection spectrogram is frequency specific.

2.4.2 Developing the Template library—Given the frequency-specific nature of spectrogram correlation, templates were first developed by finding fixed-frequency (FF) templates from a library of previously collected USVs with a mean frequency representing every 1-kHz bin from 35-kHz to 80-kHz. USVs become less likely at frequencies approaching (or exceeding) 80-kHz; therefore templates for some of the highest frequencies were not represented and templates above 80-kHz were not included. The template library was then expanded to include frequency modulated USVs (FM) and trills by using spectrogram cross-correlation to find representative USVs from a library of previously recorded USVs from our laboratory (Figure 2). To do this, clips were taken of FM USVs and trills. Sampled clips were cross-correlated in order to find USVs that exhibited a relationship with a wide variety of similar USVs (Figure 2). Clips that best represented a large group of similar USVs were then included as templates. In its final form, the optimized template library contained a total of 88 representative FF, FM, and Trill USVs.

2.4.3 Post-Detection Procedures—Two tools were developed in order to increase the efficacy of the detector. First, it was recognized that many complex USVs could be represented as the sum of a number of simpler components. For this reason, the detector was

constructed to score all instances in which spectrogram activity met the detection criteria. This method occasionally results in multiple detections of a single USV, but also allowed the detector to identify a wider variety of USVs while reducing the computational load for detection by limiting the number of templates necessary for spectrogram correlation. In order to handle redundant detections from the correlation routine, a MATLAB tool was developed which combined overlapping detections or detections which fell within 50ms of one another into a final set of individual detections.

Second, a tool was developed which allows for detections to be reviewed and for false detections to be eliminated. This tool also allows human scorers to rapidly tag USV subtypes (e.g., FF, FM or Trill). While it might seem ideal to eliminate the need for human scoring altogether, retaining a review procedure is important for a number of reasons. 1) Compared to current procedures which involve scanning entire files for USVs, the addition of a human review process to the present automated procedure comes at very little analysis 'cost'. The overall time spent scoring is greatly reduced, but rapid revision ($1.43 \pm .22$ seconds/detection) ensures data integrity. 2) False detections are inherent to any signal detection procedure. As stated above, these can be eliminated with more stringent detection thresholds. Nevertheless, conservative thresholds also increase the number of false negatives (USVs that are missed by the detector) or would require a greater library of templates and greater amount of computer processing time. Human review allows the use of less stringent detection thresholds to avoid false negatives, while insuring against false detections. 3) The detection review tool allows for visual and auditory confirmation of USVs, whereas signal detection procedures tend to work on either auditory (e.g. amplitude detectors) or visual (e.g., spectrogram correlation) representations of vocalizations.

2.4 Assessment of Detector Reliability

The number of true positives (detected USVs), false negatives (USVs that were missed by the detector), and false positives (non-USV detections) was derived by using the table comparison tool in Raven (Bioacoustics Research Program, 2011). This tool allowed for the comparison of detected USVs with those in the master table of USVs for each file. Data from all 18 recorded files were then assessed by taking the weighted detection percentages for the 18 included recordings. Weighted averaging was used in order to account for variability in the number of vocalizations (and thus putative detections) within each file. The equations for the percent detected for each file and weighted average for all files are as follows:

$$\% \text{ Detected} = \frac{\text{True Positives}}{\text{Total Number of USVs per file}}$$

$$\text{Weighted Mean} = \frac{w_1x_1 + w_2x_2 + \dots + w_nx_n}{w_1 + w_2 + \dots + w_n}$$

Where:

x_n = % detected for each file

w_n = The weight for each file (total number of USVs per file)

The percent of USVs detected was compared across call types (FF, FM, and Trills) using a series of paired-samples t-tests.

2.5 Signal-to-Noise Analysis

Signal detection problems rely on the ability of a detector to correctly isolate a signal (e.g., USVs) from background noise (e.g., ambient noise recorded by the microphone). Thus, we examined the reliability of the detector as a function of the signal-to-noise ratio of observed USVs. Signal-to-noise ratios were calculated for each USV by extracting the peak power (dB) for each call and dividing this value by the peak power (dB) of a sample of the noise-band taken from each file. Samples of the noise band were always taken from a section of the file that was free of any recording artifacts, such that they always represented ambient noise. The probability of detecting a USV based on the observed signal-to-noise ratio was then calculated for the observed range of signal-to-noise ratios.

2.6 Probability of Detection based on Call Frequency

A final analysis was detected to determine the probability of USV detection as a function of the frequency (kHz) of each detected call. An ideal detection algorithm would have a relatively equal probability of detection across the full range of 50-kHz USVs.

For analysis USVs were sampled in 5-kHz bins from 35-kHz through 70-kHz. The probability of detection was determined by calculating the number of detected USVs in each bin of frequency and dividing this number by the total number of USVs observed in that specific frequency range.

2.7 Generalizability of the detector

Recordings (n=5) were donated by Drs. Aaron Johnson, Paul Clarke, and Michelle Ciucci—all of whom have experience in USV recording and analysis (Johnson et al., 2010, 2011; Ciucci et al., 2007, 2008, 2009; Wright et al, 2010, 2011; Scardochio & Clarke, 2012). These files were used to test the generalizability of the template detector to other recording environments, recording hardware, etc. and extend results from our laboratory. These recordings were processed using identical procedures as described above. Independent-samples t-tests were then used to compare the percent of FF, FM, and Trill USVs detected for files from other laboratories against recordings taken from our laboratory.

2.8 Custom XBAT Extensions

The most important feature of the current detection procedure is the ability to create MATLAB-based extensions that can be used for custom analyses. These programs might include (but are not limited to) functions for exporting data, detecting USV subtypes, extracting salient call features, etc. Beyond the necessary extensions for concatenating detections and reviewing detections (See Section 2.3), we have implemented a number of tools that might serve to benefit the community of USV researchers. Primarily, these extensions serve to export detection data into formats commonly used by USV researchers, including, Microsoft Excel or Avisoft SASLab Pro (Avisoft Bioacoustics, Berlin, Germany).

3.0 Results

3.1 Detection Threshold

The threshold for template detection was optimized by comparing rates of false negatives and false positives across a range of correlation thresholds from 0.1 to 0.9. This procedure revealed that the detector was optimized when correlation thresholds were set between 0.2 and 0.4 (higher threshold values produced more conservative results, with fewer false detections and more false negatives, while lower thresholds produced more liberal results with fewer missed detections but an increased false positive rate). Based on this information, the threshold for USV detection and data analysis was set to 0.3 for analysis of all data described in this report.

3.2 Detection Speed

Detection using an intel CORE i7vPro 2.8GHz processor on a computer with 16gb of RAM running Windows 7 64 bit processed files, on average, at 0.43 ± 0.001 times their normal speed. In other words, each 1 second of recorded data took 2.35 ± 0.01 s to scan using the template detector. As would be expected, pilot analyses indicated that the detection speed is highly dependent on computer processing power and the number of templates included in the library for spectrogram correlation. Thus, proper hardware selection and optimization of the template library is important for the rapid detection of USVs. Nevertheless, detection speed is faster than human detection, which can take nearly ten times longer, i.e., 15-20s of human scoring time for every 1s of recorded data (Reno et al., 2013).

3.3 Detector Accuracy

Overall, the number of USVs detected corresponded highly to the total number of USVs observed in each file ($R^2=0.987$; Figure 3). Template detection in XBAT correctly identified $95.79 \pm 1.52\%$ (Weighted Mean $\pm$ Weighted SEM; 273/285) of fixed-frequency USVs, $93.79 \pm 1.16\%$ of frequency modulated USVs (1360/1450) and $90.64 \pm 2.51\%$ (779/855) of Trills (Figure 4). Paired-samples t-tests revealed that there were no significant differences in the accuracy of detection between any of these three call types [all $|t(18)| < 1.55$, $p > 0.14$, N.S.].

3.4 Probability of Detection: Signal-to-noise Analysis

Given that template detection procedures rely on correlations, it is necessary that recorded USVs are of sufficient power (dB) to be differentiable from background noise. Accordingly, results showed that the probability of detecting USVs increased as the signal-to-noise ratio (USVs : ambient noise) increased (Figure 5). Notably, when the signal-to-noise ratio was at 1.3 :1 or greater, $> 97\%$ of observed USVs were detected. On the contrary, when signal-to-noise ratios were below this threshold, USVs were poorly detected using the template detector.

Overall, this evidence suggests that human scorers are better able to score USVs with poor signal-to-noise ratios. It should not be overlooked that this is certainly due, in part, to human scorers' ability to use multiple sources of information to verify USVs (e.g., both visual and auditory confirmation). Still, by optimizing microphone placement, recording settings, and

making attempts to dampen ambient noise, it may be possible to better detect the remaining small percentage of faint calls falling within the noise band. At worst, assuming that ambient noise is random and unrelated to the animal's behavior, a slight underestimation of call frequency could occur.

3.5 Probability of Detection: Call Frequency Analysis

The probability of detection for each frequency bin across the 50-kHz range is shown in Figure 6. For reference, the total number of USVs emitted in each bin are also shown. The nomenclature for '50-kHz USVs' has been utilized because this range represents the average frequency amongst USVs in this range. Consistent with this notion, the highest number of USVs was observed in the 50-55kHz frequency bin. Notably, more than 95% of the USVs emitted in this range were detected using the template matching procedure.

The probability of detection tended to be lowest for frequencies at which very few vocalizations were observed. The probability of detection was lowest between 35- and 40-kHz (5.8%), where a total of only 17 USVs were observed. Barring the exceptionally poor detection in this range, USVs were detected with relatively equal probability across the remainder of the 50-kHz range. Nevertheless, the present analysis illustrates that the accuracy of the detector might be improved by further sampling under-represented frequencies.

3.6 Detector Generalization

The generalizability of the detector to other recording setups was tested by examining 5 files recorded from other laboratories with experience recording USVs in the rat. Independent-samples t-tests comparing of the number of FF, FM, and Trills detected for files recorded within our laboratory versus those recorded by others revealed no significant differences in accuracy when detecting fixed-frequency USVs [$t(21) = -0.76$, $p = 0.46$, N.S.], frequency-modulated USVs [$t(21) = 0.07$, $p = 0.94$, N.S.] or Trills [$t(21) = 0.51$, $p = 0.62$, N.S.]. For these files, $100 \pm 0.00\%$ of the fixed-frequency calls were detected, $94.5 \pm 8.77\%$ of frequency modulated USVs were detected and $85.17 \pm 2.65\%$ of trills were detected.

4.0 Discussion

Methods for acoustic signal detection are not novel (e.g., Brandes, 2008; Mohammad & McHugh, 2011; Chesmore & Ohya, 2004; Charif & Pitzrick, 2008; Mellinger & Clark, 1997). Still, each new species and experimental scenario presents a specific challenge for research scientists and bioacousticians and requires the development of a specific set of tools. The current detection method for 50-kHz USVs reliably extracted > 90% of recorded USVs and its performance generalized to recordings taken in other laboratories. It is also especially important to point out that vocalizations missed by the detector from all laboratories' recordings tended to be relatively quiet, (signal : noise ratios of 1.2 : 1 or lower). Thus, missed detections might be reduced, not through further changes to the detection method, but instead by optimizing the placement of recording hardware and/or making efforts to reduce ambient noise.

It should be acknowledged that human scorers are also susceptible to missed detections. Indeed, the number of missed detections made by human scorers is negatively correlated with experience (unpublished observations). This is especially important when comparing the present detection method to current standards in the field. The current detection method is highly accurate when compared against a master list of USVs and thus is even more comparable to current standards after taking into account missed detections by human scorers. Scoring by research assistants is advantageous in that humans are able to detect USVs with relatively poor signal: noise ratios and that experienced scorers are able to detect > 97% of all recorded USVs. Alternatively, template detection procedures are advantageous for their speed and relative consistency. Overall, we would suggest that the appropriate compromise between these advantages is to score subsets of files using both methods and to report statistics for inter-rater reliability between human scoring and automatic scoring while template detection procedures are still in development.

Template detection of USVs has a number of advantages over other types of detectors that might be found in the field of bioacoustics. First, template detection is superior to amplitude detectors or similar band-limited signal detection methods when used in a laboratory setting. While these detectors could prove beneficial under ideal conditions, they are less than optimal for use with operant conditioning chambers, which are conducive to high amplitude noise. Template detections are also more intuitive than ‘whistle-tracking’ procedures, which rely on extracting salient events from within the spectrogram based on changes in energy (dB) relative to background noise. Indeed, attempts by our laboratory to implement either of these methods produced highly variable results and failed to achieve the same degree of accuracy as human scorers or the current detection method. Nevertheless, it remains possible that detector accuracy or speed could be increased by combining multiple detection methods.

4.1 Limitations and Future Directions

Our primary goal in developing the present USV detector was to release it to the research community as a method/tool. Notably, the concept of template detection is not novel in the field of bioacoustics. However, the current library of templates represents an effective and tested method which is new to the community studying rat USVs. With this in mind, it is certain that the performance of the detector can be improved by adding or modifying templates over repeated testing. Indeed, the frequency specific nature of spectrogram correlation and the immense variability in the type and frequency of emitted USVs leaves ample room for improvement in the current detector as the library continues to be amended to better capture this variability. Perhaps most importantly, the open-source nature of the current method allows for the type of collaboration among USV researchers that will be necessary to implement future improvements and share novel MATLAB extensions for data analysis.

When considering the accuracy of the present detection method, one area needing improvement, in which the detector performed most poorly, was in detecting trills as compared to fixed frequency or frequency modulated call subtypes. Given the wide variety in trills emitted by rats, it is not surprising that trills were detected at slightly lower rates

than fixed-frequency or frequency-modulated USVs. It would seem that the accuracy of the detector could be improved for trills by continuing to add USVs to the template library to accommodate this variability.

Currently, the described USV detection procedure is limited to the detection of 50-kHz USVs. It is our aim to expand the detector to include USVs in the 22-kHz range. Certainly, we believe that 22-kHz USVs will present unique challenges (e.g., 22-kHz USVs are more often occluded by low-frequency artifacts), although the basic detection method would not change. Importantly, our current experience with models of drug abuse has limited the types of 22-kHz USVs that we have observed (Barker et al., 2010; 2013a; 2013b). Thus, developing an effective detector for 22-kHz USVs will require a much larger library of vocalizations which would include a variety of both short and long 22-kHz USVs. Alternatively, it should be considered that ideal detection templates could be computer generated by developing a set of commonly observed ‘syllables’ which could be iterated across the full frequency range in which USVs are emitted. This approach would be particularly advantageous for controlling the signal-to-noise or other parameters of implemented templates.

Acknowledgments

The authors thank Drs. Paul Clarke, Aaron Johnson, and Michelle Ciucci as well as Laura Grant and Tina Scardochio for gracefully providing data to test the generalizability of the detector. Second, we would like to thank Raimund Specht and the staff at Cornell's bioacoustic research program for their continued assistance. Finally, we thank Steve Simmons, Lisa Servilio, Danielle Bercovicz, Thomas Grace Sr., and Jackie Thomas, for excellent assistance. This study was supported by the National Institute on Drug Abuse Grants DA006886, DA029873 (MOW) and DA032270 (DJB) and the Aresty Center for Undergraduate Research (CH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. All coauthors have seen and approve of the contents of the manuscript.

References

- Barker DJ, Root DH, Ma S, Jha S, Megehee L, Pawlak AP, West MO. Dose-dependent differences in short ultrasonic vocalizations emitted by rats during cocaine self-administration. *Psychopharmacology*. 2010; 211(4):435–442. [PubMed: 20571780]
- Barker DJ, Bercovicz D, Servilio LC, Simmons SJ, Ma S, Root DH, West MO. Rat ultrasonic vocalizations demonstrate that the motivation to contextually reinstate cocaine-seeking behavior does not necessarily involve a hedonic response. *Addiction biology*. 2013
- Barker DJ, Simmons SJ, Servilio LC, Bercovicz D, Ma S, Root DH, West MO. Ultrasonic vocalizations: evidence for an affective opponent process during cocaine self-administration. *Psychopharmacology*. 2013:1–10.
- Bioacoustics Research Program. Raven Pro: Interactive sound analysis software (Version 1.4). Ithaca, NY: The Cornell Lab of Ornithology; 2011. Computer Software Available from <http://www.birds.cornell.edu/raven>
- Borta A, Wöhr M, Schwarting RKW. Rat ultrasonic vocalization in aversively motivated situations and the role of individual differences in anxiety-related behavior. *Behavioural brain research*. 2006; 166(2):271–280. [PubMed: 16213033]
- Brandes TS. Automated sound recording and analysis techniques for bird surveys and conservation. *Bird Conservation International*. 2008; 18(S1):S163–S173.
- Burgdorf J, Knutson B, Panksepp J, Ikemoto S. Nucleus accumbens amphetamine microinjections unconditionally elicit 50-kHz ultrasonic vocalizations in rats. *Behavioral neuroscience*. 2001; 115(4):940. [PubMed: 11508733]

- Burgdorf J, Wood PL, Kroes RA, Moskal JR, Panksepp J. Neurobiology of 50-kHz ultrasonic vocalizations in rats: electrode mapping, lesion, and pharmacology studies. *Behavioural brain research*. 2007; 182(2):274–283. [PubMed: 17449117]
- Charif RA, Pitzrick M. Automated detection of Cerulean Warbler songs using XBAT data template detector software. Preliminary Report. Bioacoustics Research Program Cornell Laboratory of Ornithology Technical Report 08-02. 2008
- Ciucci MR, Ma ST, Fox C, Kane JR, Ramig LO, Schallert T. Qualitative changes in ultrasonic vocalization in rats after unilateral dopamine depletion or haloperidol: a preliminary study. *Behavioural brain research*. 2007; 182(2):284–289. [PubMed: 17397940]
- Ciucci MR, Ma ST, Kane JR, Ahrens AM, Schallert T. Limb use and complex ultrasonic vocalization in a rat model of Parkinson's disease: Deficit-targeted training. *Parkinsonism & related disorders*. 2008; 14:S172–S175. [PubMed: 18585950]
- Ciucci MR, Ahrens AM, Ma ST, Kane JR, Windham EB, Woodlee MT, Schallert T. Reduction of dopamine synaptic activity: degradation of 50-kHz ultrasonic vocalization in rats. *Behavioral neuroscience*. 2009; 123(2):328. [PubMed: 19331456]
- Chesmore ED, Ohya E. Automated identification of field-recorded songs of four British grasshoppers using bioacoustic signal recognition. *Bulletin of Entomological Research*. 2004; 94(4):319–330. [PubMed: 15301697]
- Ciucci MR, Ma ST, Fox C, Kane JR, Ramig LO, Schallert T. Qualitative changes in ultrasonic vocalization in rats after unilateral dopamine depletion or haloperidol: a preliminary study. *Behavioural brain research*. 2007; 182(2):284–289. [PubMed: 17397940]
- Ciucci MR, Ma ST, Kane JR, Ahrens AM, Schallert T. Limb use and complex ultrasonic vocalization in a rat model of Parkinson's disease: Deficit-targeted training. *Parkinsonism & related disorders*. 2008; 14:S172–S175. [PubMed: 18585950]
- Ciucci MR, Ahrens AM, Ma ST, Kane JR, Windham EB, Woodlee MT, Schallert T. Reduction of dopamine synaptic activity: degradation of 50-kHz ultrasonic vocalization in rats. *Behavioral neuroscience*. 2009; 123(2):328. [PubMed: 19331456]
- Cuomo V, Cagiano R, De Salvia MA, Maselli MA, Renna G, Racagni G. Ultrasonic vocalization in response to unavoidable aversive stimuli in rats: effects of benzodiazepines. *Life sciences*. 1988; 43(6):485–491. [PubMed: 3136280]
- DeVry J, Benz U, Schreiber R, Traber J. Shock-induced ultrasonic vocalization in young adult rats: a model for testing putative anti-anxiety drugs. *European Journal of Pharmacology*. 1993; 249(3):331–39. [PubMed: 7904565]
- Johnson AM, Ciucci MR, Russell JA, Hammer MJ, Connor NP. Ultrasonic output from the excised rat larynx. *The Journal of the Acoustical Society of America*. 2010; 128(2):EL75–EL79. [PubMed: 20707418]
- Johnson AM, Doll EJ, Grant LM, Ringel L, Shier JN, Ciucci MR. Targeted Training of Ultrasonic Vocalizations in Aged and Parkinsonian Rats. *Journal of visualized experiments: JoVE*. 2011; (54)
- Kaltwasser TM. Acoustic startle induced ultrasonic vocalization in the rat: a novel animal model of anxiety? *Behavioural brain research*. 1991; 43(2):133–137. [PubMed: 1678264]
- Mellinger DK, Clark CW. Methods for automatic detection of mysticete sounds. *Marine & Freshwater Behaviour & Phy*. 1997; 29(1-4):163–181.
- Mohammad B, McHugh R. Automatic detection and characterization of dispersive North Atlantic right whale upcalls recorded in a shallow-water environment using a region-based active contour model. *Oceanic Engineering, IEEE Journal of*. 2011; 36(3):431–440.
- Reno JM, Marker B, Cormack LK, Schallert T, Duvauchelle CL. Automating ultrasonic vocalization analyses: The WAAVES program. *Journal of neuroscience methods*. 2013
- Scardochio T, Clarke PB. Inhibition of 50-kHz ultrasonic vocalizations by dopamine receptor subtype-selective agonists and antagonists in adult rats. *Psychopharmacology*. 2012:1–12.
- Schreiber R, Melon C, De Vry J. The role of 5-HT receptor subtypes in the anxiolytic effects of selective serotonin reuptake inhibitors in the rat ultrasonic vocalization test. *Psychopharmacology*. 1998; 135(4):383–391. [PubMed: 9539263]
- Wintink AJ, Brudzynski SM. The related roles of dopamine and glutamate in the initiation of 50-kHz ultrasonic calls in adult rats. *Pharmacology Biochemistry and Behavior*. 2001; 70(2):317–323.

- Wright JM, Gourdon JC, Clarke PB. Identification of multiple call categories within the rich repertoire of adult rat 50-kHz ultrasonic vocalizations: effects of amphetamine and social context. *Psychopharmacology*. 2010; 211(1):1–13. [PubMed: 20443111]
- Wright JM, Dobosiewicz MR, Clarke PB. α - and β -adrenergic receptors differentially modulate the emission of spontaneous and amphetamine-induced 50-kHz ultrasonic vocalizations in adult rats. *Neuropsychopharmacology*. 2011; 37(3):808–821. [PubMed: 22030713]

Highlights

- Introduces a method for the automatic detection of rat ultrasonic vocalizations.
- Automatic scoring increases speed while retaining the opportunity for human review.
- Open-sourced programming allows for collaboration among vocalization researchers.

Figure 1. An illustration of spectrogram correlation

Template spectrograms (bottom row) are compared to recorded data (middle row) using a sliding window. Amplitude values (i.e., spectrogram color scale) are compared at each frequency bin of the spectrogram at a given time lag. Correlations at each lag (top row) can range between -1 and 1. Calls that exceed an adjustable detection threshold (green) are labeled as detections. Lowering the detection threshold (yellow) will increase the sensitivity of the detector, but may also increase the rate of false detections.

Template Selection with Spectrogram Correlation

Figure 2. Spectrogram Cross-Correlation

Ideal templates are selected via spectrogram cross-correlation, which reveals clusters of similar USVs. Vocalizations in the present figure represent a cluster of similar calls, all of which exhibit the same mean frequency. Parameters of each spectrogram including call duration, signal strength, and occlusion by noise either facilitate or preclude vocalizations from serving as representative templates. Poor templates exhibit correlations above the detection threshold ($r = 0.3$) for only a small number of similar vocalizations, while vocalizations that serve as the best templates exhibit correlations with many similar USVs. Red cells indicate correlations that do not surpass the detection threshold, yellow cells indicate correlations that are above but near the threshold, and green cells indicate correlations above the correlation threshold.

Figure 3. Correlation of Detected USVs and Total USVs

The number of USVs detected by templates is shown on the X-axis, while the number of USVs in the master table for each file is shown on the Y-axis.

Figure 4. Detection Accuracy and Generalizability

Percent of fixed-frequency (FF), frequency-modulated (FM) and trill vocalizations detected using the described template matching procedure. A) Eighteen recorded files were used to test the detector on FF, FM, and Trills each with a bar showing the percent detected. Figure 4B is the same as the above graph but includes data from four files donated by other laboratories that study rat USVs to test the generalizability of the detector to other recording situations.

Figure 5. Probability of Detection as a Function of Signal Strength

The probability of detection for USVs in the present analyses as a function of observed signal-to-noise ratios. Signal strength was operationally defined as the peak power (dB) of a given vocalization, while the noise-band was defined as the average power (dB) of a portion of the spectrogram proximal to the vocalization which was free from any USVs, locomotor induced noise, etc. The probability of detection is greatly reduced when signal: noise ratios fall below 1.2 :1, whereas detection of USVs with higher signal: noise ratios are readily detected.

Figure 6. The probability of detection (black) as a function of call frequency (kHz; red) USVs were detected with relatively equal probability across nearly the entire range of 50-kHz vocalizations. Moreover, the probability of detection was highest between 50 and 55 kHz where the highest number of emitted USVs was observed. Rates of USV emission were lowest at the extremes of the 50-kHz frequency range. Detection was poorest at the lowest observed frequency ranges. These data suggest that frequencies which are poorly represented are the most difficult to accurately capture using template detection procedures.

Table 1

Starting XBAT <ul style="list-style-type: none"> Download XBAT (with USV library) from: www.rci.rutgers.edu/~markwest/ Open MATLAB Open your MATLAB path, click 'Add with Subfolders', select the downloaded folder. Click save and close Open XBAT (type XBAT at command prompt in MATLAB). Select User using dropdown menu Select Library using dropdown menu If not already created, click New under the corresponding dropdown menu Use 'Edit' to modify an existing User or Library 	Loading/Opening a Sound File <ul style="list-style-type: none"> Click New, to upload a sound file <u>Single Sound</u> – Use for single sound files <u>Filestream</u> – Use to compile all sound files in a folder into one sound <u>Folder</u> – Use to compile all sound files in a folder and its subfolders into one sound Double-click on the name of the sound file you want to open to open its sound window. The sound window features a spectrogram of the sound and a link to other menus 	Creating/Displaying Logs <ul style="list-style-type: none"> On the top menu bar of the sound window, click on 'Window' and in the drop down menu click 'Log'. This will open the Log Window, which contains all information and options for logs pertaining to the open sound. To create a new log, click New towards the top of the menu, type the name in the pop-up dialog box, and click OK The log(s) selected in the 'Display' box in the Log Window are the logs whose selections are visible in the sound window
Open/Start Detector <ul style="list-style-type: none"> On the top menu bar of the sound window, click on 'Detect' and in the drop down menu click 'Data Template', to open the Data Template Menu Click on the top tab titled 'Scan' Ensure that the Output Log is the correct one and that the threshold is set to 0.3 Click Start to begin detection A window will appear detailing time lapsed and time remaining. 	Adding/Deleting Templates <ul style="list-style-type: none"> To add a template, make a box around the portion of the spectrogram in the sound window that you want to make into a template by clicking and dragging and click Ctrl+C. To delete a template, select the template you want to delete from the drop down menu in the Data Template Window, right-click on its spectrogram and click Delete Clip 	Post-Detection Procedures <ul style="list-style-type: none"> Logs for each sound can be found in the main XBAT window For all post-detection procedures, right click on the log Click on Actions to display a menu of all available actions. The following are useful for USV Detection: <ul style="list-style-type: none"> USV Clean – Combines/ groups detected events into individual USVs Selection Review – Tool for human review of detected USVs Export – Multiple extensions for exporting detections to other programs are present