User Manual – Vibrato Scope

1. Introduction

Vibrato Scope is software for analyzing vibrato in audio recordings. Developed by Dr. Tiago Lima Bicalho Cruz, the program was created to assist researchers, educators, and vocal coaches in extracting and analyzing important vibrato parameters – such as rate, extent, jitter, Sample Entropy, and additional variability indices – to achieve a robust analysis of voice modulation.

The software operates through two main approaches:

• **Area Selection**: Allows the user to load a WAV file, view its spectrogram, and manually select regions of interest for detailed analysis.

• **Batch Process**: Enables automated analysis of multiple audio files, generating individual results for each file as well as an overall summary.

2. Overview of the Analysis Process

This section describes the complete workflow of Vibrato Scope—from audio loading to the calculation and display of vibrato parameters—including the new processing functions integrated into the workflow.

2.1. Audio Loading and Pitch Extraction

Audio Loading: The WAV file is loaded (in mono) using the *librosa* library at a standard sampling rate (for example, 22050 Hz).

2.1.1. Pitch Extraction Methods Available

VibratoScope supports four pitch extraction methods that can be selected by the user prior to running the analysis:

- Praat (Autocorrelation) – via the Parselmouth interface. Implements the robust autocorrelation algorithm used by the Praat software.
- YIN (pyin) – provided through the Librosa library, using a frame size of 2048 and hop size of 512 samples.
- Harmonic Product Spectrum (HPS) – internally implemented using short-time Fourier transform (STFT), computing the product of decimated spectra to find the fundamental frequency.
- REAPER – an external tool integrated via command-line, known for its robust pitch-tracking in expressive voice signals.

These methods vary in accuracy and computational cost, and the user may select the most appropriate option based on the analysis context.

2.12. Harmonic Fallback Routine

If no valid vibrato cycles are detected after pitch extraction (for example, in low-SNR segments or due to detection failure), the software triggers a **harmonic fallback routine**:

- A short-time Fourier transform (STFT) is computed on the signal.
- The strongest harmonic above 50 Hz is selected.
- The fundamental frequency is estimated by dividing this partial's frequency by its most likely harmonic number (closest to the running median of the original pitch contour).

This ensures that the software provides usable pitch estimates even in challenging acoustic conditions.

2.2. Conversion to Cents and Centering

• **Conversion to Cents**: Each pitch value is converted to the cents scale using the formula:

$$\text{cents} = 1200 \times \log_2 \left(\frac{\text{pitch (Hz)}}{440} \right)$$

where 440 Hz is the reference (A4).

• **Centering the Signal**: The cents signal is centered by subtracting the median of the valid values. This centering removes the overall trend of the signal, highlighting the variations that correspond to vibrato.

Note: Immediately after conversion, the outlier filtering function (*filter_pitch_outliers*) is applied to replace highly discrepant values with *NaN*, ensuring a more robust basis for subsequent calculations.

2.3. Resampling to a Uniform Grid

• **Objective**: Resampling the signal to a fixed rate ensures that data are collected at regular intervals. For example, by resampling at 100 Hz, the signal is measured every 0.01 second (1/100), standardizing the time series and enabling precise comparisons in parameter calculations – such as the application of the band-pass filter and the analysis of vibrato cycles. Note that 100 Hz is only an example and this rate may be adjusted according to the needs of the analysis.

2.4. Band-Pass Filtering

• 3 to 9 Hz Filter: A fourth-order Butterworth band-pass filter is applied to the centered and resampled cents signal to isolate the vibrato oscillations, which generally occur in the 3 to 9 Hz range.

• **Why 3–9 Hz?** This range is typical for vibrato oscillations in the voice, discarding slower variations (such as melodic changes or inflections) and faster variations (noise or artifacts unrelated to vibrato).

• I**mpact**: The filtering "cleans" the signal, concentrating the algorithm on the periodic oscillations essential for vibrato. This step also prepares the data for advanced calculations such as obtaining the Sample Entropy.

2.5. Detection of Vibrato Cycles

• **Detection of Peaks and Troughs**: With the filtered signal, the software detects peaks (maximum points) and troughs (minimum points), which represent changes in the direction of the oscillation.

• **Cycle Definition**: Each consecutive pair of a peak and a trough (or vice versa) defines a "half cycle" of vibrato.

• Vibrato Rate (Rate): Calculated as:

$$\mathrm{Rate} = \frac{1}{2 \times \text{half-cycle time}}$$

This formula considers that a complete vibrato cycle consists of two half cycles.

• **Vibrato Extent (Extent):** Determined as half the variation (in cents) between the peak and the trough, representing the amplitude of the vibrato oscillations.

2.6. Smoothing of Parameters

• **Moving Average (Smoothing):** To reduce the influence of fluctuations or "noise" in the signal, a moving average is applied to the parameters of each cycle.

• **Default Window of 3 Cycles:** The default value is 3, meaning that for each cycle, the smoothed value is the average of the central cycle and its two neighbors.

▪ **Values Less Than 3:** Minimal smoothing, resulting in a noisier signal with more point-to-point variation.

▪ **Values Greater Than 3:** Excessive smoothing, which may "erase" important vibrato details or introduce a delay in representing the oscillations.

2.7. Calculation of Vibrato Metrics

After detecting and smoothing the cycles, the software calculates several metrics:

• **Average Vibrato Rate:** The mean of the rates calculated for each half cycle, converted to Hz.

• **Average Vibrato Extent:** The average of the amplitudes (in cents) of each half cycle.

• **Jitter:** Measures the variability of the cycle durations, calculated as the ratio of the standard deviation to the mean of the cycle times.

• **Coefficient of Variation (CV):** Calculated for both rate and extent using the formula:

$$\mathrm{CV} = \left(\frac{\text{Standard Deviation}}{\text{Mean}}\right) \times 100\%$$

This value indicates the relative dispersion of the data.

2.8. Conversion of Frequency to Musical Note

• **Function:** *frequency_to_note_name(freq, ref_hz=440.0)*

• **Description:** This function converts a frequency value (in Hz) to the corresponding musical note name, using A4 (440 Hz) as the reference. It is based on the MIDI system, in which each note is assigned a corresponding number (for instance, A4 is represented by 69). The function calculates a theoretical note number for a given frequency and, upon rounding, determines the specific note (C, C#, D, etc.) and its octave. For example:

• If freq = 440 Hz, the formula yields 69, which identifies the note A4.

• If freq = 523.25 Hz (approximately the frequency of C5), the rounded result is about 72, corresponding to C5.

• Formula Used:

$$\text{note\_number} = 12 \cdot \log_2\left(\frac{\text{freq}}{\text{ref\_hz}}\right) + 69$$

Here, the expression calculates how many semitones the frequency is away from 440 Hz, and adding 69 gives the MIDI note number. This value is then rounded to identify the exact note and its octave.

2.9. Outlier Filtering for Pitch

• **Function:** *filter_pitch_outliers(pitch_values, threshold=3)*

• **Description:** Removes pitch values that are considered outliers. The function uses the median of the valid values and the Median Absolute Deviation (MAD) to determine an acceptable range. Values that deviate more than the multiple defined by the threshold are replaced with NaN.

• **Benefits:** Ensures that subsequent analyses are performed on more reliable pitch data by eliminating spurious values.

2.10. Calculation of Sample Entropy

• **Function:** *sample_entropy(time_series, m=2, r=None)*

• **Description:** Calculates the Sample Entropy of a time series – a measure of the complexity or irregularity of the signal.

• **General Formula:**

$$\text{Sample Entropy} = -\ln\left(\frac{\phi(m+1)}{\phi(m)}\right)$$

where:

– *φ(m)* represents the probability (or relative frequency) that two segments of length *m* are similar (i.e., all their samples lie within a tolerance *r*).

– *φ(m+1)* does the same for segments of length *m+1*.

• **Parameters:**

– *m*: Length of the pattern (typically 2).

– *r*: Tolerance for comparing patterns. If not specified, 0.2 times the standard deviation of the series is used.

• **Application:** Useful for quantifying the regularity of vibrato data (for example, the rates and extents of cycles).

• **Interpretation:** Lower values indicate a more regular signal; higher values suggest greater complexity.

2.11. Calculation of Jitter and Coefficient of Variation (CV) Metrics

• **Jitter Function**: *compute_jitter_metrics(periods)*

Calculates multiple jitter indices (local, RAP, PPQ5, DDP) from the vibrato cycle periods, expressed as percentages or in milliseconds.

• **CV Function:** compute_cv(cycle_times, cycle_extents)

Determines the coefficient of variation for both vibrato rate and extent by computing the ratio of the standard deviation to the mean (multiplied by 100).

• **Utility:** These indices help in understanding the regularity of vibrato cycles, providing comparative measures across different recordings or analysis regions.

3. How to Use Vibrato Scope

3.1. Area Selection Mode

1. **Load the WAV File:** Click “Select WAV File” and choose the desired audio file.

2. **View the Spectrogram:** The audio spectrogram is displayed, allowing visualization of the sound components.

3. **Select Regions of Interest:** Use the mouse to click and drag over the spectrogram to define the areas where vibrato will be analyzed. The selected regions are listed on the interface.

4. **Run the Analysis:** Click “Run Analysis” to have the software:

- Extract and process the pitch (including outlier filtering);
- Convert the values to cents and center the signal;
- Resample the data and apply the band-pass filter;
- Detect vibrato cycles and calculate metrics (rate, extent, jitter, CV, Sample Entropy, etc.);
- Optionally computes the harmonic fallback estimate in segments with no detected vibrato cycles, ensuring continuous output for difficult input signals.

Automatically generate illustrative figures (comparing the signal before/after filtering, displaying peaks and troughs, and vibrato rate graphs).

5. **Saving the Results:** At the end of processing, by selecting the destination folder, the results (detailed, summary, and regional CSV files, as well as PNG figures) are saved automatically – eliminating the need to click “Save Results.”

3.2. Batch Process Mode

1. **Select Multiple Files:** Click “Batch Process” and select the WAV files you wish to analyze in batch.

2. **Define the Destination Folder:** Choose the folder where the batch process results will be saved.

3. **Automatic Processing:** For each file, the software performs all steps – extraction, filtering, cycle detection, and metric calculation – ignoring regions with fewer than 3 vibrato cycles to ensure data reliability. Individual CSV files (detailed, summary, and

regional) and generated figures (before/after filtering, peak/trough detection, vibrato rate graphs) are saved automatically.

4. **Completion:** Upon finishing, a message confirms that processing is complete and the results are available in the selected folder.

5. **Figures:**

• Figure 1 – Before and After Filtering: Generates two comparative graphs showing the cents signal before filtering and the filtered signal (3–9 Hz).

• Figure 2 – Peak and Trough Detection: Visually highlights the peaks and troughs in the filtered signal, illustrating the vibrato oscillations.

• Figure 3 – Vibrato Rate per Half Cycle: Displays the evolution of the vibrato rate (raw and smoothed) over the detected half cycles.

• Figure 4 – Summary PNG Panel: A six-panel graphical report is generated for each region or file, containing:

- Raw and smoothed vibrato rate (Hz),
- Raw and smoothed vibrato extent (cents),
- Normalized coefficient of variation (CoV),
- Bar graphs of Sample Entropy (SampEn) for rate, extent, and cycle duration,
- A summary table of global values (pitch range, jitter indices, SampEn, CoV, etc.).

Note: For batch processing or when multiple regions are selected in single-file mode, one figure is generated per file (batch process) or per selected region (single-file mode).

4. Technical Concepts and Parameters

4.1. Band-Pass Filtering (3–9 Hz)

• **Purpose:** Band-pass filtering is applied to isolate the vibrato oscillations (typically between 3 and 9 Hz) from the rest of the signal. This step removes slow components (trends, melodic inflections) and high-frequency noise, focusing the analysis on the periodic oscillations that define vibrato.

• **Effect on the Signal:** Although filtering alters the signal, it does not compromise the relevant data because the 3–9 Hz range exactly corresponds to the expected frequency of voice vibration. Therefore, the vibrato rate is calculated based solely on the cycles present in this range.

4.2. Smoothing – 3-Cycle Window

• **Definition:** Smoothing is applied using a moving average that computes the average of each cycle's parameters by considering the central cycle and its two neighbors (a 3-cycle window).

• Why Use a Default Value of 3?

  • **Balance:** A value of 3 offers a good balance between smoothing out short-term fluctuations (noise) and preserving the true vibrato characteristics.

  • **Lower Values (e.g., 1 or 2):** Minimal smoothing, which may result in a noisier signal with variations that do not represent the true vibrato.

  • **Higher Values (e.g., 5 or 7):** Excessive smoothing, which may erase important vibrato details or introduce a delay or distortion in representing the oscillations.

4.3. Calculation of Vibrato Parameters

• Vibrato Rate: Calculated as:

$$\text{Rate} = \frac{1}{2 \times \text{half-cycle time}}$$

Explanation: Each half cycle of vibrato (from a peak to a trough) represents half of a complete oscillatory cycle. Thus, the average half-cycle time is used to calculate the total oscillation rate.

• **Vibrato Extent:** Determined as half the variation (in cents) between peaks and troughs.

Explanation: This measure reflects the amplitude or "depth" of the vibrato oscillations.

• **Jitter and Coefficient of Variation (CV):**

– **Jitter:** Calculated as the ratio of the standard deviation of cycle times to their mean, indicating the regularity of the vibrato cycles.

– **CV:** Calculated for both rate and extent using the formula:

$$\text{CV} = \left(\frac{\text{Standard Deviation}}{\text{Mean}} \right) \times 100\%$$

Explanation: The CV provides a relative measure of the dispersion of the values, allowing for comparison of variability between different recordings or regions.

5. Final Considerations

- **Pitch Extraction Flexibility**: VibratoScope allows users to switch between four pitch extraction algorithms—Praat, YIN, HPS, and REAPER—depending on their needs and the properties of the audio being analyzed. This flexibility enhances compatibility with diverse recording conditions and research paradigms.

- **Fallback and Robustness**: Even when vibrato cycles are not directly detected, the system activates a fallback mechanism to estimate pitch via harmonic analysis, improving robustness in suboptimal audio recordings
- **Personalization:** Parameters such as the filter range (3–9 Hz) and the smoothing window (default of 3 cycles) can be adjusted to suit the specific characteristics of the audio being analyzed or user preferences.
- **Results:** The outcomes are generated as CSV files (detailed, summary, and regional) and PNG figures, allowing the user to independently visualize and analyze the data.
- **Responsible Use:** The software is intended for personal and research use. Although the calculations are precise, the results are approximate and do not replace critical analysis or expert judgment.

6. Support and Contact

If you have any questions, issues, or suggestions, please contact:

Dr. Tiago Lima Bicalho Cruz

Email: tiagolbc@gmail.com

Telephone: +55 31 99113-9509