The Journal of Open Source Software

VibratoScope: A Python Toolkit for High-Resolution Vibrato Analysis in Singing Voice

Tiago Lima Bicalho Cruz [1]

1 Federal University of Minas Gerais (UFMG), Brazil

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Fabian-Robert Stöter

Reviewers:

- @ax-le
- @hadware

Submitted: 26 May 2025
Published: unpublished

Summary

Vocal vibrato is a defining feature of trained singing, particularly in Western classical music, where it is described as a periodic fluctuation in the fundamental frequency (f_0), often accompanied by coordinated variations in loudness and timbre. Foundational definitions [Seashore (1932); Sundberg (1994)] characterize good vibrato as a pulsation that enhances tonal richness and expressiveness.

Its acoustic analysis typically includes parameters such as vibrato rate (Hz), extent (cents or semitones), and regularity. The latter is often assessed via jitter and shimmer, representing cycle-to-cycle variations in frequency and amplitude, respectively. More advanced descriptors like sample entropy, determinism, and line length have been used to quantify vibrato complexity and stability [Manfredi et al. (2015); Acosta Martínez & Daffern (2023)].

[Manfredi et al. (2015)] proposed a high-resolution method using the BioVoice software, referencing extent norms reported. [Ferrante (2011); Anand et al. (2012); Capobianco et al. (2023)] highlighted stylistic differences, showing that Early Music singing features faster, narrower, and less regular vibrato compared to Romantic style. Variability across genres and historical contexts has also been observed in jazz [Manfredi et al. (2015)], operetta and schlager [Nestorova et al. (2023)], and contemporary commercial music (CCM) [Hakanpää et al. (2021)].

Additionally, [Glasner & Johnson (2022)] noted that historical recording technology may have influenced vibrato perception and performance in modern opera singers. These findings reinforce the view of vibrato as a stylistically dependent and context-sensitive vocal parameter.

Installation

To install VibratoScope, ensure Python 3.9 or higher is installed. Then, execute the following commands:

```
git clone https://github.com/tiagolbc/vibratoscope.git
cd vibratoscope
pip install -r requirements.txt
python vibratoscope.py
```

The repository includes a requirements.txt file listing all dependencies, ensuring a straightforward setup across platforms.

Example Use

VibratoScope supports both interactive and automated workflows:

Interactive Analysis:

- Launch the GUI by running `python vibratoscope.py`.
- Load a .wav file and select a region of interest using the GUI's time-domain viewer.
- Click "Run Analysis" to compute vibrato metrics, which are displayed as plots and saved as CSV/PNG files.

Batch Processing:

- Select multiple .wav files via the GUI or command-line interface.
- Run the analysis without manual region selection, using default or user-specified parameters.
- Results are organized in structured folders, with each file generating corresponding CSV and PNG outputs.

These workflows make VibratoScope versatile for both detailed case studies and large-scale dataset analysis.

Validation and Testing

VibratoScope includes a set of pre-analyzed audio files and outputs in the `examples/` directory. These synthetic test cases contain singing vowel sounds with known vibrato parameters (e.g., 5.0 Hz rate, 0.3 semitone extent) and are used to validate the vibrato detection algorithms.

Each example provides: - A `.wav` file with controlled vibrato features - Output figures including pitch traces, cycle-by-cycle plots, entropy, and summary analysis - Corresponding CSV files with extracted metrics

These test cases confirm that VibratoScope reliably measures vibrato parameters across known input conditions, validating its use for both research and pedagogical applications.

Figure **??** illustrates the analysis of a synthetic vowel with 5.0 Hz vibrato rate and 0.3 semitone extent, showing accurate pitch detection, vibrato cycle extraction, and summary visualization.

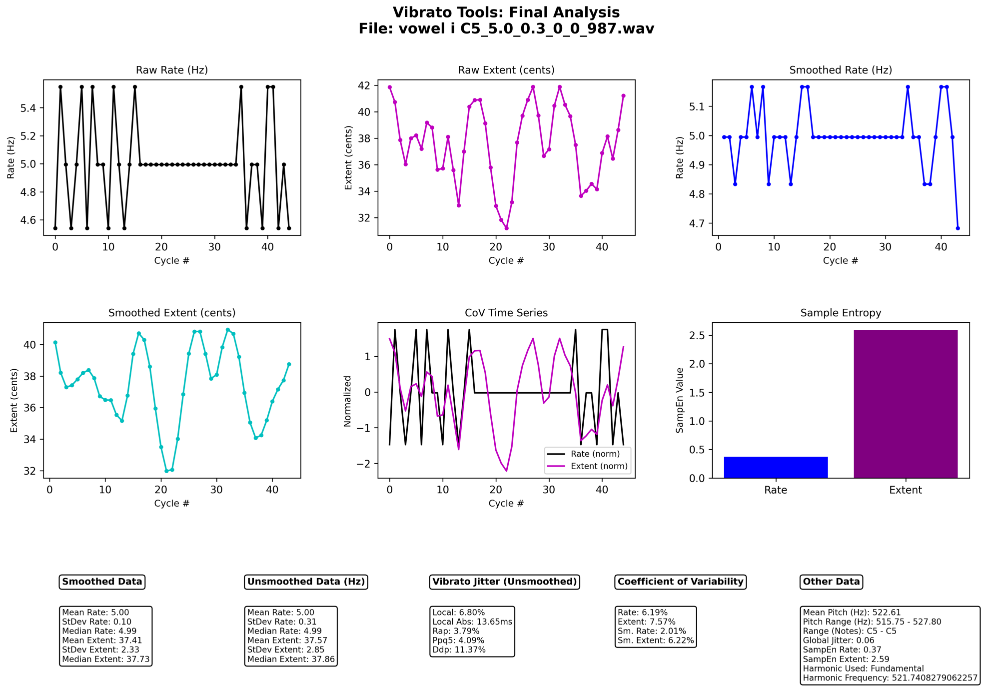

Figure 1: Example output of synthetic vowel test

Statement of Need

Vibrato is a cornerstone of vocal expression, yet its quantitative analysis is often constrained by the limitations of existing tools. Proprietary software like VoceVista Pro provides real-time vibrato overlays on spectrograms but is hindered by opaque algorithms, limited export options, and commercial licensing costs, reducing transparency and reproducibility [VoceVista (2022)]. BioVoice offers high-resolution estimates of vibrato rate, extent, jitter, and shimmer, but its Windows-only executable lacks batch processing and an API, limiting integration with automated pipelines [Morelli et al. (2019)]. Open-source alternatives, such as the Embodied Music Lab (EML) Vibrato Tools [Howell & Nestorova (2025)], are cross-platform Praat plugins that provide a user-friendly graphical interface and do not require coding skills or manual scripting. These tools offer default configurations that work for most cases. While the need to install Praat and the plugin may represent an additional step, the EML tools are highly accessible and well-documented. The Vibrato Analysis Toolbox (VAT) provides a sophisticated Hilbert-transform pipeline with user-definable filters, yet its reliance on MATLAB ties it to costly licenses and demands signal-processing expertise, restricting its accessibility [Zhang et al. (2017)].

VibratoScope was designed to address these gaps, offering a comprehensive, open-source solution for high-throughput vibrato analysis. Implemented in Python and released under an MIT license, it combines: (i) an intuitive graphical user interface (GUI) for interactive audio selection, pitch extraction, and vibrato cycle detection, accessible to non-programmers; (ii) a command-line interface supporting unattended batch processing with multiple pitch extraction methods, including Praat [Boersma (1993)], YIN [Cheveigné & Kawahara (2002)], Harmonic Product Spectrum [Noll (1970)], and REAPER [Talkin (2015)]; and (iii) transparent CSV and PNG outputs that integrate seamlessly with statistical and machine-learning workflows. By merging the ease-of-use of commercial GUI software, the extensibility of script-based tools, and the accessibility of open-source code, VibratoScope enables reproducible research for voice scientists, pedagogues, and clinicians. Its class-based API allows developers to extend functionality, while its platform-independent design supports diverse applications, from vocal pedagogy to cross-cultural studies of singing styles. Adopted in academic settings, including graduate courses and workshops at conferences like the Pan-European Voice Conference, VibratoScope meets the growing demand for open, scalable tools in voice research.

Mathematics

Vibrato is modeled as a periodic modulation of the fundamental frequency (f_0(t)). The vibrato extent (in cents) is calculated as:

$$\text{Extent} = 1200 \cdot \log_2 \left(\frac{f_0(t)}{f_{\text{mean}}} \right)$$

where (f_0(t)) is the instantaneous fundamental frequency, and (f_{mean}) is the mean frequency over the analyzed segment. The vibrato rate is derived from the frequency of this modulation, typically computed via Fourier analysis or autocorrelation.

For example, the sample entropy ((SampEn)) quantifies the irregularity of vibrato oscillations:

$$\text{SampEn}(m, r, N) = -\ln \left(\frac{A}{B} \right)$$

where (m) is the pattern length, (r) is the tolerance, (N) is the number of data points, and (A) and (B) are counts of matching patterns within the time series. These metrics are implemented in VibratoScope to provide robust descriptors of vibrato behavior.

Figures

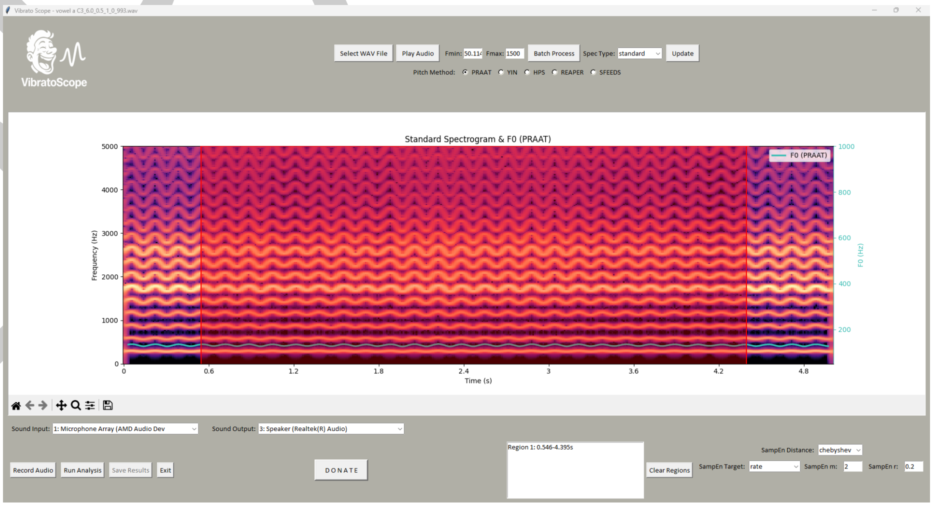

Figure 2: Graphical User Interface

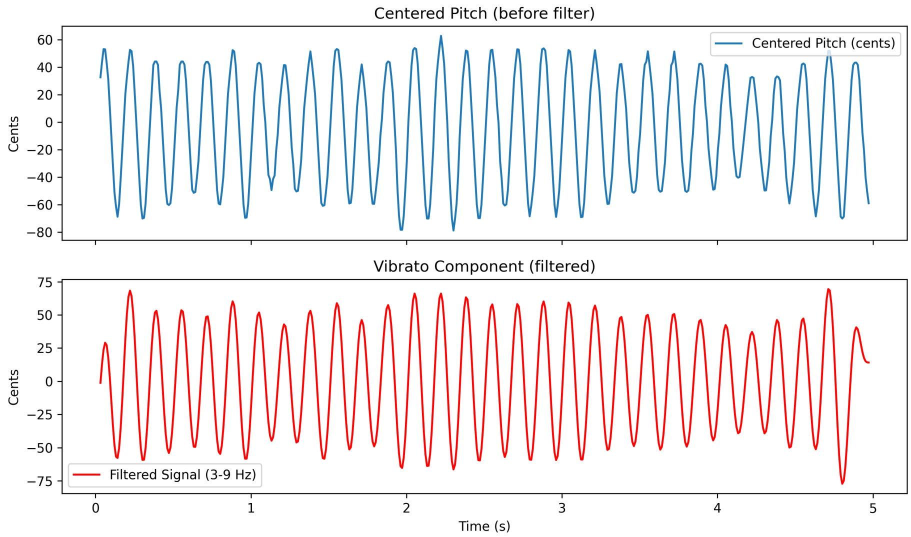

Figure 3: Pitch Filtering

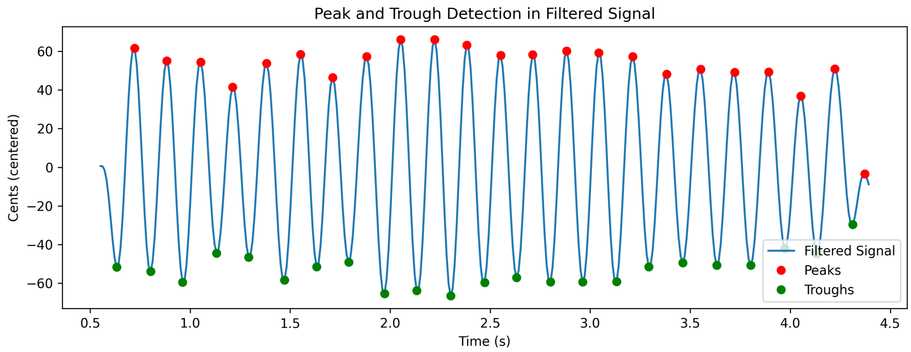

Figure 4: Peak and Trough Detection

The Journal of Open Source Software

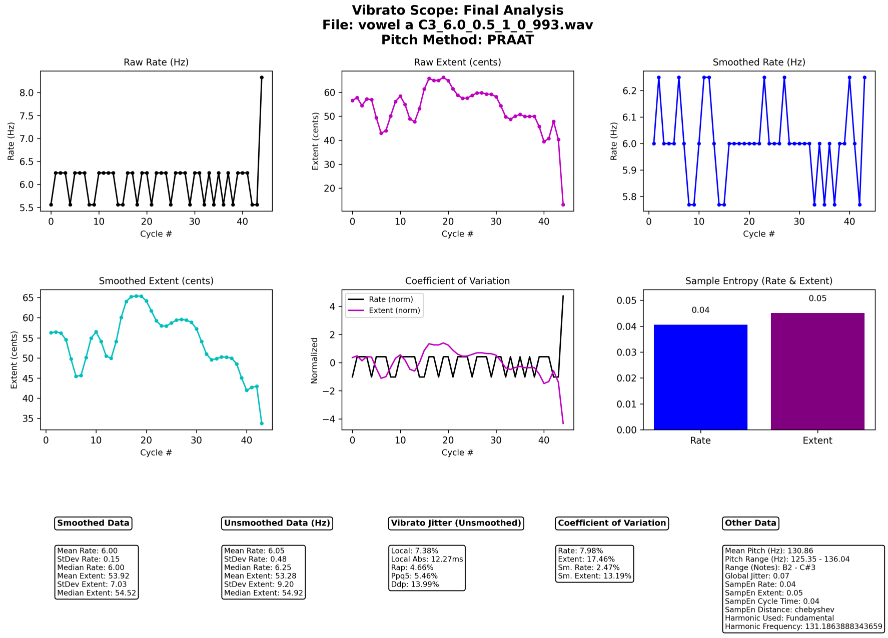

Figure 5: Final Analysis Summary

Acknowledgements

VibratoScope was developed as part of doctoral research at the Federal University of Minas Gerais (UFMG).

References

Acosta Martínez, G., & Daffern, H. (2023). Complexity of vocal vibrato in opera and jazz recordings: Insights from entropy and recurrence analyses. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2023.11.020

Anand, S., Wingate, J. M., Smith, B., & Shrivastav, R. (2012). Acoustic parameters critical for an appropriate vibrato. *Journal of Voice*, *26*(6), 820.e19–820.e25. https://doi.org/10.1016/j.jvoice.2012.06.004

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proceedings of the Institute of Phonetic Sciences*, *17*, 97–110.

Capobianco, S., Nacci, A., Calcinoni, O., Bruschini, L., Berrettini, S., & Bottalico, P. (2023). Assessing acoustic parameters in early music and romantic operatic singing. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2023.02.009

Cheveigné, A. de, & Kawahara, H. (2002). YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, *111*(4), 1917–1930. https://doi.org/10.1121/1.1458024

Ferrante, I. (2011). Vibrato rate and extent in soprano voice: A survey on one century of singing. *The Journal of the Acoustical Society of America*, *130*(3), 1683–1688. https://doi.org/10.1121/1.3613930

Glasner, J. D., & Johnson, A. M. (2022). Effects of historical recording technology on vibrato in modern-day opera singers. *Journal of Voice*, *36*(4), 464–478. https://doi.org/10.1016/j.

jvoice.2020.07.022

Hakanpää, M., Tolonen, T., & Tervaniemi, M. (2021). Acoustic features in classical and commercial music singers: Vibrato and beyond. *Frontiers in Psychology*, *12*, 682935. https://doi.org/10.3389/fpsyg.2021.682935

Howell, I., & Nestorova, T. (2025). *EML Vibrato Tools*. https://www.embodiedmusiclab.com/eml-tools-sign-up.

Manfredi, C., Barbagallo, D., Baracca, G., Orlandi, S., Bandini, A., & Dejonckere, P. H. (2015). Automatic assessment of acoustic parameters of the singing voice: Application to professional western operatic and jazz singers. *Journal of Voice*, *29*(4), 517.e1–517.e9. https://doi.org/10.1016/j.jvoice.2014.09.014

Morelli, M., D'Errico, G., & Manfredi, C. (2019). BioVoice: High-resolution vibrato analysis software for clinical and musical applications. *SoftwareX*, *9*, 275–281. https://doi.org/10.1016/j.softx.2019.02.007

Nestorova, T., Brandner, M., Gingras, B., & Herbst, C. T. (2023). Vocal vibrato characteristics in historical and contemporary opera, operetta, and schlager. *Journal of Voice*. https://doi.org/10.1016/j.jvoice.2022.12.027

Noll, A. M. (1970). Pitch determination of human speech by the harmonic product spectrum, the harmonic sum spectrum and a maximum likelihood estimate. *The Journal of the Acoustical Society of America*, *47*(2B), 667–685. https://doi.org/10.1121/1.1911939

Seashore, C. E. (1932). *The vibrato*. University of Iowa Press.

Sundberg, J. (1994). Acoustic and psychoacoustic aspects of vocal vibrato. *STL-QPSR*, *35*(2–3), 45–67.

Talkin, D. (2015). *REAPER: Robust epoch and pitch estimator*. https://github.com/google/REAPER.

VoceVista. (2022). *VoceVista video pro – user guide and documentation*. "https://www.vocevista.com/".

Zhang, Y., Drioli, C., & De Poli, G. (2017). A MATLAB toolbox for vibrato analysis in singing. *Computer Music Journal*, *41*(1), 45–60. https://doi.org/10.1162/COMJ_a_00419