Teach Yourself Georgian Folk Songs Dataset: An Annotated Corpus of Traditional Vocal Polyphony

David Gillman[1], Atalay Kutlay[1], Uday Goyat[2]

[1]New College of Florida, Sarasota FL, USA.
[2]Georgia Institute of Technology, Atlanta GA, USA.

Our Goals

- To construct a challenging dataset for algorithms for multiple pitch estimation.
- To help ethnomusicologists study traditional polyphonic songs from the Republic of Georgia.
- To help singers learn traditional Georgian songs.

What's Here

- A DATASET of traditional vocal polyphony annotated with an estimate of the fundamental frequency f_0 ("pitch") of each vocal part, and a measure of confidence in the estimate.
- AN INTERACTIVE WEB VISUALIZATION which plays the audio and displays the vocal parts.
- A PROCESS AND WORKFLOW for annotating the dataset, which uses multiple recordings of each song and takes advantage of the salience of individual voices in each recording.

Georgian Polyphony

Traditional Georgian three-part vocal harmony varies widely from region to region.

Audio Sources

Two folk song collections were recorded in 2004 by master musicians of Guria and Samegrelo.

There are one recording of the mix and three recordings of the parts. The part recordings are polyphonic, but one voice is at higher volume.

Process and Workflow Step 1. Monophonic Pitch Estimates

Input: One song, consisting of one recording of each part, with other parts audible at lower volume.

Preprocessing: Voiced/unvoiced detection

Processing: Six pitch estimates using techniques for monophonic audio:

- Fourier analysis [1, 2]
- Autocorellation [3, 4]
- Wavelet analysis [5]
- Deep learning [6]

Postprocessing:

- Alignment of estimates
- Note estimation, as an aid for the analyst

Step 2. Interactive Tool

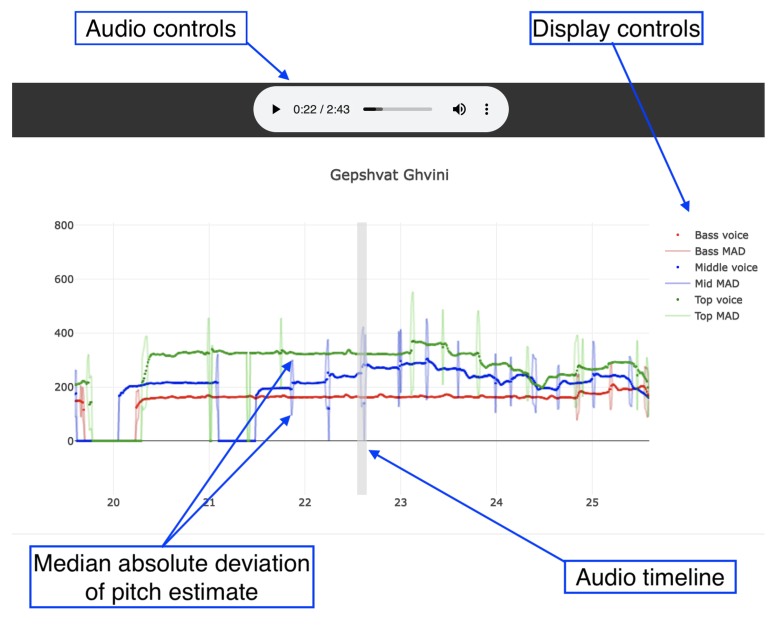
The analyst uses a graphical interface to construct a target pitch estimate for each part by stitching together the best estimates for different sections of the part. In sections of a part where no estimate is correct the analyst can specify a pitch range.

Main functionalities of the tool.

Step 3. Median of Re-estimates

- Rerun monophonic pitch estimates, constrained:
  - within manually set pitch ranges
  - within 10% of target pitch
- Final pitch estimate is median of estimates
- Confidence interval is median absolute deviation

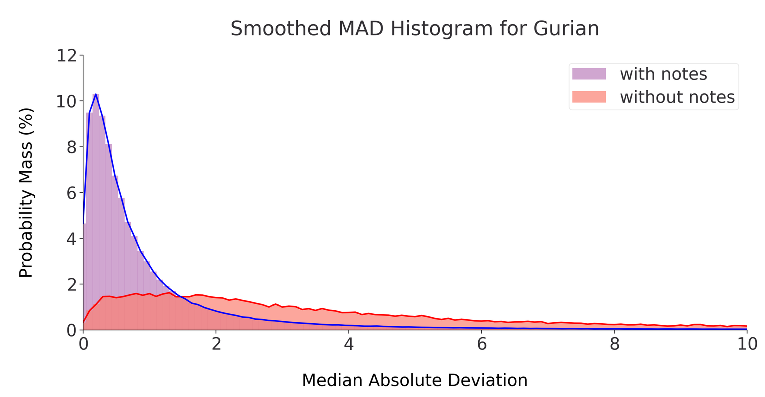
Interactive Web Visualization of Annotated Dataset



Audio-visual display of the song *Gepshvat Ghvini*

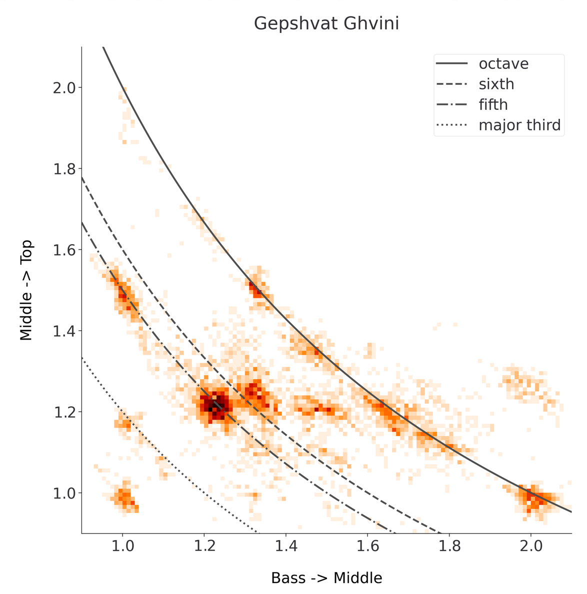
Confidence in Method

Our confidence in the final estimate is the variability of the monophonic estimates, measured as median absolute deviation. This variability is smaller during identified notes than during transitions between notes, as expected



Insights into Georgian Music

This heatmap histogram shows the chords of one song. The triadic chords cluster around those with a "neutral" third between the minor and major third.



X-axis: ratio of pitches, middle:bass. Y-axis: ratio of pitches, top:middle. The dark patch at $\sim (1.22,\ 1.22)$ indicates a neutral third.

Further Work

- Apply the method to new song collections.
- Use the data to improve multipitch estimation.
- Develop the web visualization
  - as a learning tool for Georgian singers.
  - as a tool for ethnomusicologists.

Contact information:
David Gillman
dgillman@ncf.edu

References

[1] A. Michael Noll. Short-Time Spectrum and "Cepstrum" Techniques for Vocal-Pitch Detection. *J. Acoustical Soc. Am.*, 36(296), 1964. doi: 10.1121/1.1918949.
[2] Dik Hermes. Measurement of pitch by subharmonic summation. *The Journal of the Acoustical Society of America*, 83(257), 1988. doi: 10.1121/1.396427.
[3] Paul Boersma. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In *IFA Proceedings 17*, pages 97–110, 1993.
[4] A. de Cheveigné and H. Kawahara. YIN, a fundamental frequency estimator for speech and music. *J. Acoustical Soc. Am.*, 111(4), 2002. doi: 10.1121/1.1458024.
[5] E. Larson and Ross K. Maddox. Real-Time Time-Domain Pitch Tracking Using Wavelets. In *Proc. Univ. of Illinois at Urbana Champaign REU Program*, 2005.
[6] J. W. Kim, J. Salamon, P. Li, and J. Bello. Crepe: A Convolutional Representation for Pitch Estimation. *2018 IEEE ICASSP*, 2018. doi: 10.1109/ICASSP.2018.8461329.