VIBRATO ANALYSIS DATASET AND DERIVATIVE TRAJECTORY PARAMETERIZATION

Doga Buse Cavdir
Stanford University
cavdir@stanford.edu

Guiseppe Bandiera, Oriol Romani
Universitat Pompeu Fabra
guiseppe.bandiera@upf.edu
oriol.romani@upf.edu

Xavier Serra
Universitat Pompeu Fabra
xavier.serra@upf.edu

ABSTRACT

Vibrato is one of the most widely studied musical features over years. From singing voice to numerous instruments, vibrato shows both very characteristic and also distinctive behavior within instruments. However, there is no detailed vibrato dataset available to MIR research to the best of our knowledge. In this research, we present an extendable vibrato dataset including musical techniques and musical expressions. We further present the ongoing automatic detection algorithm and derivative trajectory analysis of pitch contours. In addition to algorithm's easiness in implementation and short run-time duration, the quantitative evaluation shows success rate of 89.7% in detecting the vibrato presence among 173 vibrato sounds and 112 no-vibrato sounds. The parameters extracted from the derivative trajectories are also included with the sound files and source code to have a reproducible analysis. With prior instrument detection; it will offer an enhanced analysis, detection method and an enlarged dataset.

1. INTRODUCTION

The wide range of usage of vibrato brings problems as unpredictable behavior, distinct characteristics depending on the instruments and error-prone analysis resulting from applying the Fourier transform to the f_0 of the audio [1,4]. This extended abstract presents a vibrato dataset of five instruments with distinct musical techniques and a methodology for the detection of vibrato presence based on the first derivation of the pitch contours. The data is collected and features are extracted as a part of the previous research [2]. In this current research, the performance of the algorithm is enhanced, the extend and the functionality of the dataset are introduced. The bassoon subset is included and ground truth annotations for subsets and overall dataset are provided within the new release of the dataset. Furthermore, the preliminary instrument detection for vibrato detection and extension of the dataset in the context of instruments, performers and musical techniques are proposed as a future work. On the way to producing reproducible research, recordings, annotations, statistical parameters of the derivative trajectory and the source code are given in the dataset.

2. DATASET

2.1 Contribution to MIR

The idea of creating this dataset derived from the challenges of reaching data for automatic vibrato detection. A collection of vibrato tones at different pitch ranges with their corresponding no vibrato tones is limited in the field. Moreover, existing datasets are missing musical techniques which create a great distinction in vibrato behavior. Even when vibrato is applied for the same note, the finger position and on which note it is applied affect the width, extend and the rate profoundly [4,5]. With this motivation, we aimed to provide a vibrato dataset regarding the expressiveness of performers to MIR community. As a first step, we extended the dataset by adding bassoon recordings, which introduced a low pitch range instrument vibrato, and including the ground truth files. A portion of the dataset is devoted to monophonic melodies and articulations including scales, staccato and legato playing. With these additional features of rate, note and melody forms, instrument range and quality, the dataset allow MIR researchers to use our data in musical expressiveness detection and performance evaluation as well as it provides a great utility and resource for automatic detection.

2.2 Data Collection and Content

The recordings of all three sub-datasets are made in Phonos Recording Studio in MTG. Transverse flute, violin and violoncello recordings are collected from another dataset released as a part of former MTG researcher Carlos Vaquero Master's thesis [3]. The availability of reaching vibrato tones and their corresponding non-vibrato tones was one of the main reasons in selecting the dataset. Another important reason is that it offers many more instrument in a broader pitch range which is useful in extending the research as the future work. The second dataset created in the Good-sounds context broadens the research in terms of vibrato quality, melody, scale and triads as well as articulations, staccato and legato techniques. The biggest contribution to the dataset is from the note and melody forms at three different rate specially recorded for this research again in Phonos, MTG. Over all, the biggest reason why those subsets are selected is that they are open-access resources and curated in MTG. The number of sounds in each subset and features if subsets can be seen in Table 1-2 respectively.

2.3 Feature Extraction

The feature extraction from derivative trajectories are applied to vibrato and no vibrato pairs. Having different rates of vibrato allowed us to compare vibrato behavior under different musical expressions of the performer. Pitch contours are obtained using Pitch Contours Melody algorithm from Essentia standard library. Peak value rate is observed to be significantly affected by the rate of vibrato.

Dataset	Note-Melody	# of vibrato sounds	# of no vibrato sounds	# of total sounds
Carlos Vaquero	note	86	83	169
Good-sounds	note	10	-	51
	melody	32	9	
MTG	note	87	29	124
	melody	4	4	

Table 1. Vibrato-non-vibrato pairs and note-melody distribution among the subsets.

Dataset	Instru-ments	Pitch range	Additional Features
Carlos Vaquero	Flute, violin, violoncello bassoon	C4-G6, G4-G5, D3-G4, C2-C4	• Vibrato - no vibrato pairs • Pitch range • Open source
Good-sounds	alto-sax	-	• Quality tags • Articulations and melody
MTG	violin	G3-B5	•Vibrato rates • Vibrato- no vibrato pairs • Melody

Table 2. Features of subsets

This observation led us to extract features as peak percentage, mean peak difference, maximum peak difference, max difference location. The parameters for each subset note sounds are annotated and provided in the dataset. Additionally, environmental noise reduction function based on the pitch confidence is created as a part of this research and provided in the dataset.

2.4 Evaluation and Results

The extracted features in the previous research are used to develop the algorithm for each instrument in this current research. Quantitative evaluation of the proposed method is presented. The current performance rate in detecting vibrato is measured as 89.9% and 91.67% in violin recordings in MTG and Carlos Vaquero subsets respectively where this rate was 72.9% for previous f_0 detection method applied to our violin recordings [6]. For violoncello, the rate dropped to 75% and for transverse flute to 59.4% due to environment noise in non-vibrato sounds. As an outcome of the current research, it is observed that the algorithm performance and success rate in detection is significantly increased when the vibrato sounds are preprocessed by supervised instrument detection. The success rate in detection for transverse flute is increased to 79.7%. As a next step, the algorithm will be enhanced by prior parameter estimation according to detected instrument.

3. CONCLUSION AND FUTURE WORK

In this research, we proposed vibrato dataset and the evaluation of derivative based analysis. The dataset includes different musical techniques applied with vibrato as articulations, staccato and legato. Another important aspect of the dataset is that it provides different lengths and rates of vibrato depending on varying finger position techniques [5]. Our approach in the analysis of dataset is a first step in the automatic vibrato detection algorithm for a variety of instruments. Previous vibrato researches in the MIR field are limited to mid pitch range string instruments. However, as a next step, prior instrument detection is being implemented. This way, we aim to apply the proposed methodology to a wider instrument range and to expand the dataset. The dataset includes significant amount of musical techniques highly linked with the performance styles and skills. Another additional contribution to the dataset as a future work will be comparing different performers on the same technique. Finally, by including sounds, source code where the detection function and noise elimination proposed and the annotations of each parameters, vibrato durations, and rate, we aimed to create an algorithm and a dataset reproducible by others.

4. REFERENCES

[1] L. Yang and E. Chew, "Ava: An Interactive System for Visual and Quantitative Analyses of Vibrato and Portamento Performance Styles," *International Conference on Music Information Retrieval*, 2016.

[2] D. Cavdir, G. Bandiera, O. Romani and X. Serra. "Vibrato Analysis Dataset, Detection and Parametrization", Zenodo, 2016.

[3] C. Vaquero. "Improving the description of instrumental orchestral sounds through ontologies and automatic content analysis in the context of Freesound," Master's thesis, Universitat Pompeu Fabra, 2012.

[4] C. Hsu and J. R. Jang, "Singing pitch extraction by voice vibrato/tremolo estimation and instrument partial deletion," *International Conference on Music Information Retrieval*, pp. 525–530, 2010.

[5] R. Timmers, and P. Desain. "Vibrato: Questions and Answers from Musicians and Science," *Proceedings of the sixth ICMPC*. Keele, 2000.

[6] P. Herrera, & J. Bonada. (1998). "Vibrato Extraction and Parameterization in the Spectral Modeling Synthesis framework." DAFx, 1998.