Pitch Detection for Music in Noisy Environments

Performance evaluation of BaNa: a Hybrid Approach for Noise Resilient Pitch Detection

Myron Vasilik, Logan Stillings, and Carmen Cortazar
Final Project
ECE 472 - Audio Signal Processing
University of Rochester, Rochester, NY

Abstract

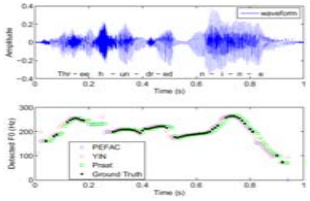
A noise resilient pitch detection algorithm, BaNa, is evaluated to test its accuracy and effectiveness in achieving the lowest Gross Pitch Error (GPE) rate in noisy environments. Pitch detection in noisy environments remains at the forefront of current research in music signal analysis with emerging applications such as speech perception and recognition music notation programs. An analysis of BaNa is provided and will mimic the results obtained in the BaNa paper. Furthermore, additional iterations will be performed changing various parameters such as, timestep and frequency bounds of the signal . Lastly, the results obtained in the experiment were compared to those presented in the BaNa paper. The results show that BaNa provides lower GPE rates than other pitch detection algorithms in music evaluation of noise resilient pitch detection. BaNa achieves the lowest GPE rate with the most consistency and is responsive to parameter changes.

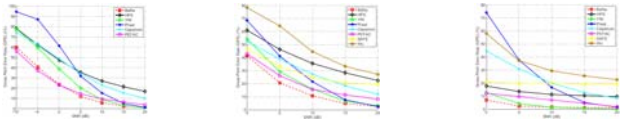
Related Work

BaNa Fo Detection Algorithm for Speech

1. Preprocessing
2. Determination of Fo candidates
3. Selection of Fo from candidates

Experimental Settings:
-Speech and noise databases
-Error measurement metric
-Parameter Tuning

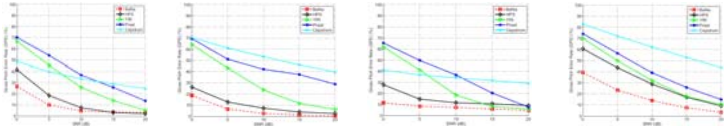




Figures are the GPE rate of the different algorithms for the LDC database

BaNa Fo Detection Algorithm for Speech and Music

1. Modification of the BaNa code for speech
2. Experimental settings for Fo detection for music
3. Parameter Tuning
4. BaNa vs BaNa music



Figures are the GPE rates of different algorithms for Violin, Trumpet, Clarinet & Piano

Research Proposal

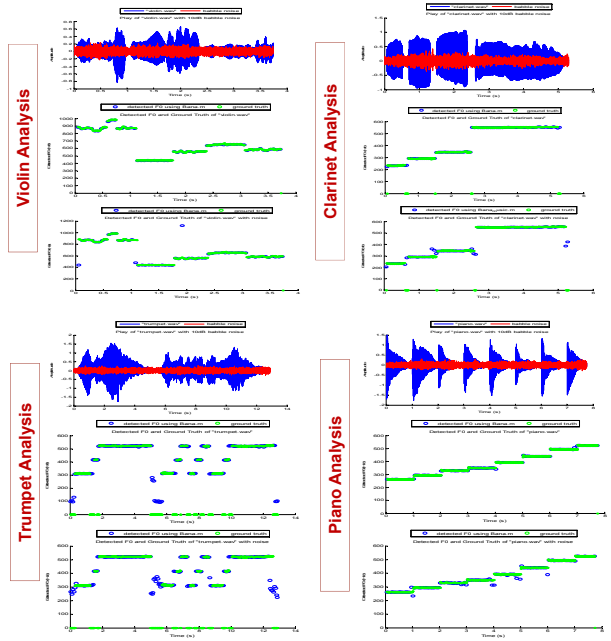
❑ **Mimic the results presented in the BaNa paper**
"Bana: A Noise Resilient Fundamental Frequency Detection Algorithm for speech and music"

❑ **Sound Samples**
to include:
1.) pre-recorded sound samples from BaNa project archives
2.) sound samples recorded using the same instruments used in BaNa project
3.) sound samples using instrument not used in BaNa (i.e. guitar and marimba)

❑ **Evaluate BaNa and its effective to achieve the lowest GPE in music signals**

Experiment & Results

Experiment - Part A: Mimic results in BaNa paper [2]



Results – Part A

❑ **BaNa_music.m**

Violin: - most accurate predictions when n=0 and n>0
- discrepancies found are gaps between notes

Clarinet: - estimated Fo nearly perfect when noise=0
- estimated Fo is slightly off when noise>0: due to noise in gaps where the clarinetist changes notes

Trumpet: - Fo predictions lose accuracy when n>0
- when n=0 clumps present at 100 Hz
- when n>0 clumps present at 300Hz
*algorithm may regard the frequencies of noise as the Fo

Piano: - Fo when noise=0 is very accurate
- predictions deviate when noise>0
- discrepancies in noise signal at end of notes: due to decaying sound
*algorithm may regard the frequencies of noise as the Fo

Experiment & Results

Experiment – Part B: Evaluation of BaNa – using sound samples recorded for this project

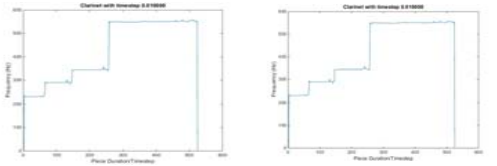
❑ **BaNa_music_auto.m**

Clarinet: - Bana_music Fo detection algorithm finds the Fo in each frame
- BaNa archived 'Clarinet.wav' file is a *midi* file

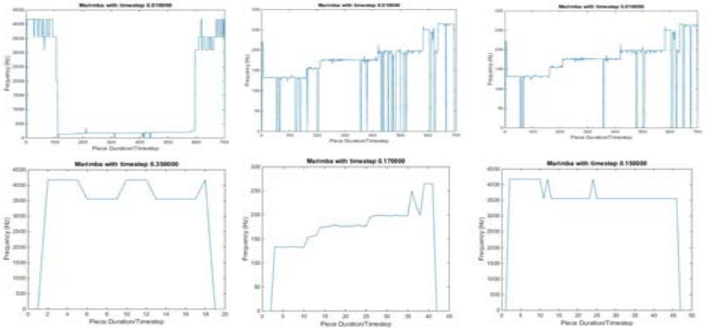
❑ **Live recoding of instrument—Marimba**

- BaNa detects high-frequency recording noise not apparent to listener and labels it as Fo content
- **Since Fo detection is the purpose of this project:**
Find a way to accurately detect to Fo information of the marimba recording when n=0 and n>0

Clarinet Analysis



Marimba Analysis



Results – Part B

Timestep is very unpredictable—a difference as small as .02s can change the algorithm's perception of the signal entirely; however, narrowing the frequency bounds down to an acceptable range, this issue would subside. This is seen in the figures above. We used the default Fomin= 20Hz and Fomax= 4200Hz and changing only the timestep.

Solution: (1) Evaluate signal at default frequency bounds, varying timestep
(2) Obtain detected max/min Fo bounds of signal (discard anomalous values)
(3) Use detected max/min values as new bounds

References

[1] N. Yang:, H. Ba, W. Cai, W. Heinzelman and I. Demirkol. "BaNa: A Noise Resilient Fundamental Frequency Detection Algorithm for Speech and Music," *IEEE/ACM Transactions on Audio, Speech, and Language Processing 22.*

[2] H. Ba, N. Yang, I. Demirkol and W. Heinzelman, "BaNa: A Hybrid Approach for Noise Resilient Pitch Detection," *Proceedings of the 2012 IEEE Statistical Signal Processing Workshop (SSP '12), Aug. 2012.*

[3] J.P Bello, G. Monti, and M. Sandler, "Techniques for automatic music transcription," in *Intl Symposium on Music Information Retrieval, 2000, pp. 23-25.*