Sounding the Keys: A Comprehensive FFT-Based Approach for Piano Note Detection

Neethu Marriam Joji[1], Pooja Namjoshi[2], Jerryn Joeboy[3], Suraj Khandare[4]

[1]Student, Dept. of Information Technology, Fr. C. Rodrigues Institute of Technology, Navi Mumbai
[2]Student, Dept. of Information Technology, Fr. C. Rodrigues Institute of Technology, Navi Mumbai
[3]Student, Dept. of Information Technology, Fr. C. Rodrigues Institute of Technology, Navi Mumbai
[4]Professor, Dept. of Information Technology, Fr. C. Rodrigues Institute of Technology, Navi Mumbai

---

Abstract – *Piano - a popular skill for both young and adults alike, is something that requires proper practice and guidance. When learning a new skill, a tutor plays an important role as the learner familiarizes themselves with the foreign skill. However, a tutor cannot always monitor and guide the student. While feedback from a professional educator is highly valuable, this traditional method of learning the piano is not viable for all; appointing a tutor becomes tedious and expensive. While practice makes perfect, it is equally important to practice correctly. Lack of proper guidance at the early stages of learning may lead to the development of incorrect practices. Timely and consistent feedback motivates players and ensures accurate practices from early learning stages. This research explores the application of signal processing, specifically utilizing the Fast Fourier Transform algorithm. By leveraging FFT's capabilities, our research demonstrates the successful development of models that exhibit accurate note identification. The findings finally highlight the transformative potential of integrating signal processing into the piano learning process. This research aims to make learning the arts more accessible to all, to ensure that keen learners are not hindered from learning the piano. The proposed system is designed to assess piano performances in real-time using the Fast Fourier Transform (FFT) algorithm and thresholding techniques for note identification. The proposed algorithm aims to improve the music evaluation process by providing real-time assessment, eventually enhancing the quality of automated assessments of piano performances. This paper's main emphasis is on the implementation of the FFT algorithm, thresholding methods, and real-time feedback.*

Key Words: **Fast Fourier Transfer algorithm, High Frequency Content (HFC) detection, thresholding methods, and real-time assessment.**

1.INTRODUCTION

A piano is an instrument that produces sound by striking strings with hammers. Pianos have a wide frequency range, from about 27 Hz to 4186 Hz. Learning a piano and mastering it takes many years and a lot of practice. Practicing with proper guidance every time is not feasible for many aspiring pianists. Students look to seek feedback on their performances in order to improve their technique and musicality. Traditional methods of assessing piano playing, such as hiring a music teacher can be time consuming and economically taxing. To address these challenges, we propose a solution that implements digital signal processing techniques to provide real-time feedback and performance assessment for piano players. The Fast Fourier Transform (FFT) algorithm is used for note detection and thresholding techniques are used to separate noise from the musical notes. The developed algorithm takes piano audio samples as input and generates a real-time 'correctness' score to verify how accurately the piano piece was played.

The paper focuses on the implementation of signal pre-processing and signal analysis functions to identify piano notes and assess them against their defined sheet music. Our algorithm aims to overcome the gas in existing systems by improving the accuracy of note detection and implementing real-time evaluation.

2.PROPOSED SYSTEM

The proposed algorithm aims to automate the process of students receiving feedback for their piano play. The algorithm, written in Python, accepts audio input in MP3 format (.mp3 and .m4a). The algorithm's main features include audio pre-processing to reduce noise and assess audio accuracy. The raw audio received consists of noise along with the play. In order to extract the clean audio, the algorithm thus employs a pre-processing algorithm to reduce the noise and assess the audio's accuracy.

Using the Fast Fourier Transform algorithm, the proposed system assesses the processed audio file. Factors such as frequency of the note, and rhythm of the piano piece are evaluated to calculate a correctness score for the piano play. The algorithm assesses the piano play against the predefined sheet music.

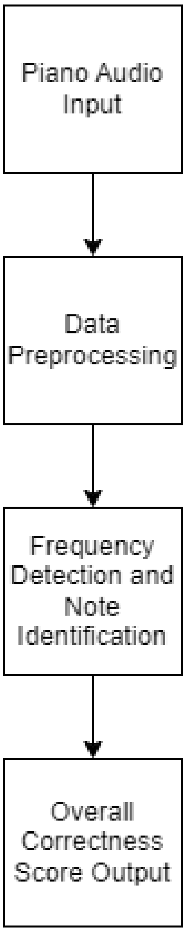

Fig. 1: Block diagram

2.1 Hann Windowing

Hann windowing, or Hanning window, is used for the audio input's pre-processing. The audio signal is segmented into frames of 20 milliseconds. The Hann function is applied to each frame. The processed audio segments are used, in tandem with the raw audio segments to minimize abrupt transitions and reduce noise. The individual segments are then combined to form the complete audio with reduced noise and undesired jumps in the audio. The Hann windowing facilitates further spectral analysis of the audio file, increasing the efficiency of the proposed algorithm in assessing the frequency of the piano notes played.

2.2 Fast Fourier Transform

The Fast Fourier Transform algorithm is used to convert the audio signal from a time domain signal to a frequency domain signal and to further perform spectral analysis to identify the piano notes. FFT breaks down a signal into its constituent frequencies, revealing the amplitudes and phases of the frequencies. FFT is a more efficient way of computing the Discrete Fourier Transform (DFT). Due to the improved efficiency of FFT algorithm, it is better suited for the real time analysis of the piano notes; which our proposed system aims to achieve. The audio signal is segmented into frames of 20 milliseconds, to accurately represent the pitch of a musical note in the small timeframe. The FFT function is applied to each frame to convert the time domain segments to frequency domain signals. The function analyses each frame to identify the dominant frequencies. As pianos keys have unique frequencies, by identifying the dominant frequency in the frame, we can subsequently identify the note played in the piano sample.

2.3 High-Frequency Content Detection

Onset detection is the process of identifying the beginning of all individual notes in audio samples. The proposed algorithm utilizes the high-frequency content (HFC) function to identify the individual piano notes in the piano play. The frequency-domain audio segments are inputted to the HFC functions. The HFC onset detection function detects the high-frequency content and identifies the peaks as the onsets of the piano notes. Piano notes are characteristically high-frequency notes. The frequency peaks are matched with the defined dataset of piano notes' frequencies. Thus, the use of the HFC onset detection function allows for accurate detection of the piano notes of the audio sample.

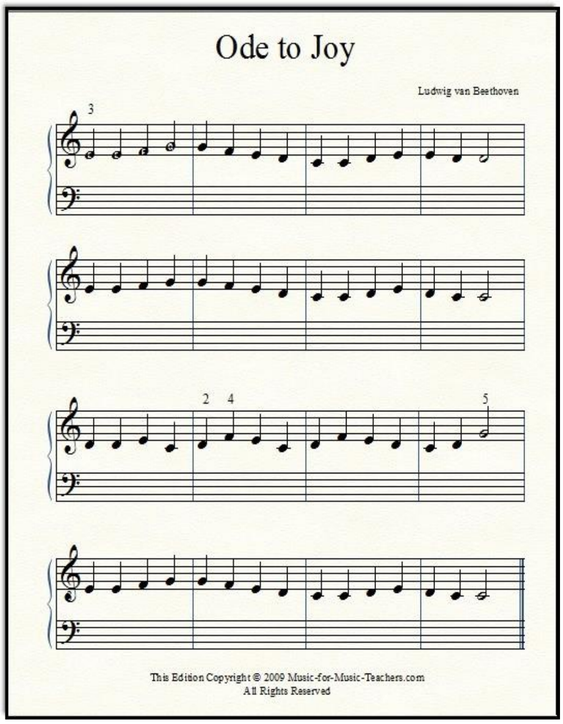

Fig. 2: Sheet Music for Ode to Joy [8]

In the figure 3, the x-axis represents the time in millisecond to showcase the segments of the 'Ode to Joy' audio samples. The y-axis represents the amplitudes of the audio signal. The HFC function can identify the onset of the notes in the Ode to Joy audio sample based on the frequencies of the notes defined in the sheet music, as per figure 2.

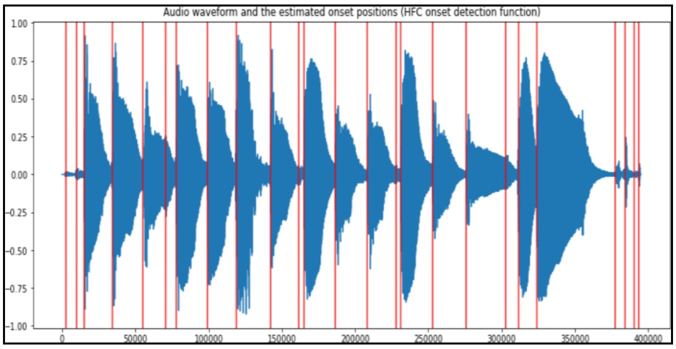

Fig. 3: HFC graph showing the onset detection of notes

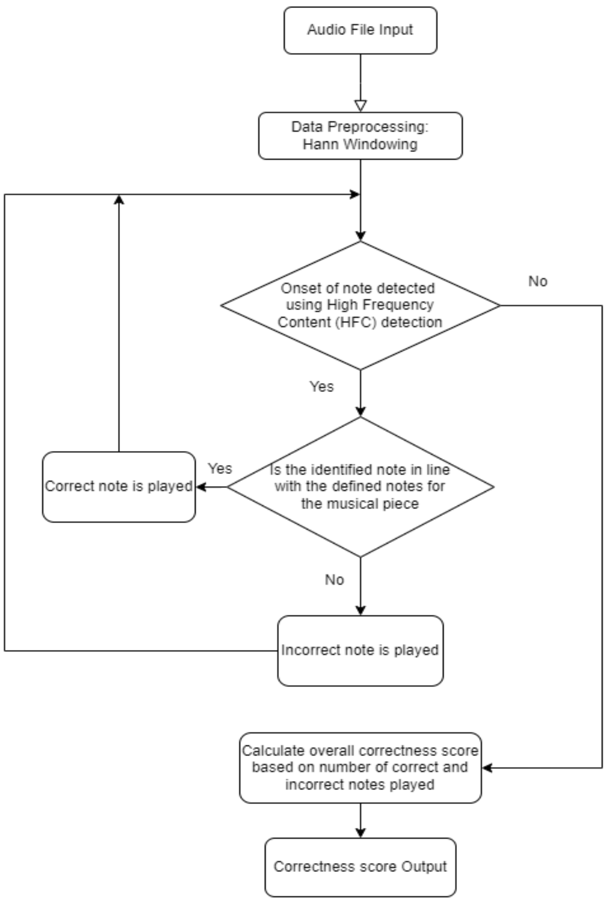

Fig. 4: Flow diagram

3. RESULT

In this study, the algorithm's performance was evaluated using the piano piece 'Ode to Joy,' which consists of 60 notes. Three separate plays, executed by the same player, are labeled as Test Sample 1, Test Sample 2, and Test Sample 3, and the correct play of the piece. The proposed system was tested against the assessment of a professional wherein the parameters used were the number of notes played correctly, the number of notes missed, incorrect notes played, and the overall correctness score. The correct audio sample when assessed by the system provided a correctness score of 98. The professional gave a score of 100 to the control audio sample, leading to a difference of 2 between the accuracy scores for the sample audio file. The difference in the overall correctness score calculated by the proposed system and the professional was 2, 0, and 1 respectively for the 3 test sample audios.

Table -1: Human vs. Machine Audio Sample Assessment – Correct Audio Sample

	Correct Audio Sample	
	Assessed by Professional	Assessed by Proposed system
Number of notes played correct	60	59
Number of notes missed	0	1
Incorrect notes played	0	0
Overall correctness score	100	98

Table -2: Human vs. Machine Audio Sample Assessment – Audio Sample 1

	Test Sample 1	
	Assessed by Professional	Assessed by Proposed system
Number of notes played correct	15	14
Number of notes missed	20	16
Incorrect notes played	25	30
Overall correctness score	25	23

Table -3: Human vs. Machine Audio Sample Assessment – Audio Sample 2

	Test Sample 2	
	Assessed by Professional	Assessed by Proposed system
Number of notes played correct	33	33
Number of notes missed	9	10
Incorrect notes played	18	17
Overall correctness score	55	55

Table -4: Human vs. Machine Audio Sample Assessment – Audio Sample 3

	Test Sample 3	
	Assessed by Professional	Assessed by Proposed system
Number of notes played correct	55	56
Number of notes missed	2	0
Incorrect notes played	3	4
Overall correctness score	92	93

4. CONCLUSIONS

By utilizing the Fast Fourier Transform for audio signal processing, we have implemented the conversion of an audio signal from its time domain to the frequency domain. Conversion of the audio signal in the frequency domain has allowed us to identify individual piano notes and how accurate they are to the original piano piece. While the proposed algorithm has achieved an accuracy rate of 94% in identifying the piano notes, improvements can be made to accommodate complex piano plays and nuances of the players' musicality. Future works may focus on the assessment of the audio quality and timbre of the piano play to provide a more holistic assessment.

REFERENCES

[1] J.K. Patel and E.S. Gopi," Musical Notes Identification Using Digital Signal Processing", Procedia Computer Science, vol. 57, pp. 876-884, 3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015)

[2] Sumarno, Linggo & Iswanjono, Iswanjono. (2017). Feature Extraction of Musical Instrument Tones using FFT and Segment Averaging. Telkomnika (Telecommunication Computing Electronics and Control).15.1280-1289. 10.12928/TELKOMNIKA.v15i3.3381.

[3] S.W. Foo and P.L. Wong," Recognition of piano notes" in Proceedings of the International Conference on Information, Communications and Signal Processing, 1999: National University of Singapore

[4] Zielinski, Tomasz. (2021). Fast Fourier Transform.

[5] Settel, J. and Lippe, Cort. (1994). Real-time musical applications using the FFT-based resynthesis.

[6] Monnier, Nolan & Ghali, Darien & Liu, Sophie. (2021). FFT and Machine Learning Application on Major Chord Recognition. 426-429. 10.1109/ICUFN49451.2021.9528762.

[7] Podder, Prajoy & Khan, Tanvir & Khan, Mamdudul & Rahman, M. (2014). Comparative Performance Analysis of Hamming, Hanning and Blackman Window. International Journal of Computer Applications. 96. 1-7. 10.5120/16891-6927.

[8] https://www.music-for-music-teachers.com/ode-to-joy.html