Music Notes Recogniser using Digital Signal Processing

2EC502- Digital signal Processing

Lab Project

Electronics and Communication Engineering department

Institute of Technology, Nirma University

Submitted By:-

Mann Raval — 19BEC109

Maarjani Sanghavi — 19BEC117

Semester-5

Batch-B2

Submitted To:-

Dr. Bhupendra Fataniya

Department of Electronics and Communication Engineering,

Institute of Technology, Nirma University.

Introduction

Extracting relevant musical information from a live or recorded performance is simple for an experienced listener, but it is far more challenging for a learner or machine. For a variety of practical applications, it would be ideal to collect this information in a quick, error-free, automated manner. The development of a software system is the subject of this thesis. that takes a digitised waveform representing an input as input attempts to extract notes from an acoustic music signal. In order to create a musical score, you must first receive a signal. Detecting events, or precisely where the individual notes are located within the signal start and stop, as well as pitch extraction, or the recognition of owing to the pitches. In both the temporal and frequency domains, careful signal processing is required. The problem is solved via domain signals. This project's main purpose is to provide a learning aid. Musicians, producers, composers, DJs, remixers, and educators will benefit from this as well as music students This project might be described as a box that takes any type of music as input and outputs the song's lyrics features. The purpose of this study is to come up with techniques for the process of analysing and characterising a signal in order to determine musical parameters can be easily and objectively retrieved a recurring issue. The process for obtaining such is described in the musical literature. In terms of signal, this is especially true.

Literature survey

A. Sound: The three parameters below can be used to characterise a sound. Pitch is the first thing that comes to mind when thinking of sound.

• Pitch

• Quality

• Loadness

The human ear perceives the frequency of a sound as pitch. A note with a high frequency has a high pitch, while a note with a low frequency has a low pitch. The pitch is low when the frequency is low. A pure tone is a sound that is not distorted in any way. Tuning fork or electrical signal generator is only one frequency in this. Due of its higher intensity. The fundamental note has the greatest amplitude and is the most noticeable the majority The other frequencies are called overtones or harmonics 2f0, 3f0, 3f0, and so on are sound quality codes. The sensation of being excessively loud is based on physiological factors.

B. Musical notes: Humans can hear signals with frequencies between 20 and 20 kHz. Some of this wide range is taken into consideration. Piano is linked to it. Different pianos come in a wide range of price ranges. Each piano note has its own personality. As seen in Figure 1, A note like C represents the basic frequency. D,...etc. The 12 half steps in the final C are separated into two halves. Distance from the one before it, with a fundamental frequency of is two times as high As a result, this portion (from one C)One has been removed. The interval between C and the next C is known as an octave. C1, C2, and C3 are the letters of the alphabet such that distinct octaves can be distinguished.

C. Equation of frequency.

The basic formula of the notes of the tempered scale is given by

$$f_n = f_0 \times (a)^n$$

Here, f0 the frequency of a single fixed note that needs to be determined.

D. Sampling

A sound wave is an analogue wave of changing air pressure produced by your voice (or a musical instrument). To store a sound wave, a computer, on the other hand, must record discrete values at discrete time intervals. Therefore it is necessary to convert the analogue signal into discrete values. Sampling is a technique for capturing discrete temporal values. The process of recording distinct pressures is known as quantizing. In recording studios, the standard sampling frequency is 48 kHz. CDs have a 44.1 KHz sampling rate. The bare minimum for sampling the signal's frequency should be at least twice that of the maximum the audio's frequency. Humans can hear frequencies between 20 and 20 KHz, which explains why sampling frequencies in the 40 KHz range are so popular. Sampling is done at higher rate so that none of the information is lost and the quantization noise is reduced.

$$fs \geq fmax$$

E. Frequency and Fourier Transforms

The index corresponding to the highest amplitude represents the most significant frequency component, as shown in fig 2, which may be calculated using the method.

$$f = (i / T) \times fs$$

Where i = Index at which the maximum amplitude exists, T = Total samples of FFT at a time, fs is the sampling frequency.

F. Padding with zeros

Despite the fact that padding with zeroes is a typical and effective strategy in many applications, it does not appear to improve our data in this case. The frequency resolution is doubled, but the data isn't any better. While the peaks are rounder due to higher frequency resolution, the difference between the highest point on the original sample and the highest point on the padded sample is just 1 Hz at most, indicating that padding with a lot more zeroes is probably not worth the effort, especially at low frequencies. The purpose of padding zeros is to increase the size of the input sequence by a factor of two. To make the total number of samples equal to the next higher power of two, padding of zeros is done.

Methodology

The audio signal file is made with different frequency components. Those frequencies are obtained through Fast Fourier Transform Algorithm. The audio signal is preprocessed by Averaging, Thresholding, Selecting the window operations. Then the audio signal is padded with zeros, computation through Fast Fourier Transform Algorithm and lastly the assignments of the frequencies takes place to the respective notes. This whole process is divided into detection and identification which is summarised into a flowchart. In time domain, the signal is visible as per waveform.

Flowchart

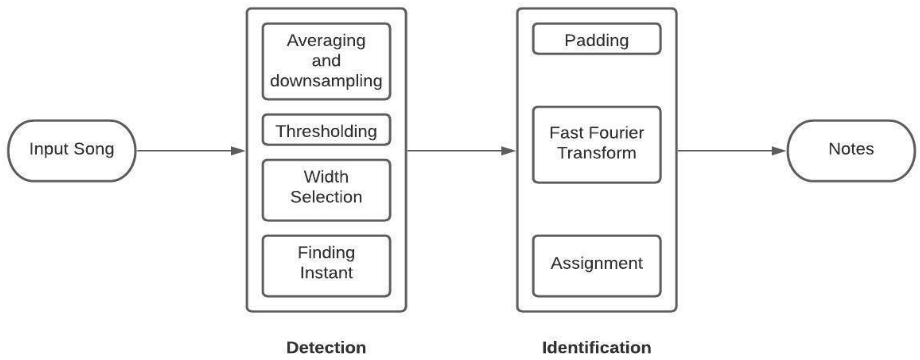


Fig.1 Flowchart

Simulation

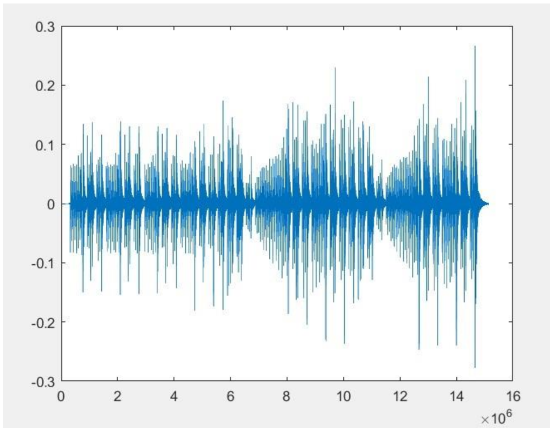


Fig. 2. Time Domain representation of Audio Signal

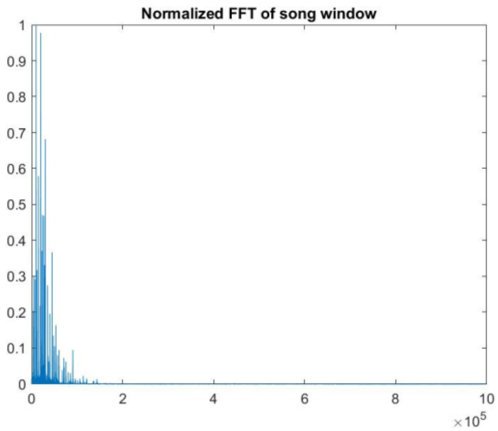


Fig 3. FFT of the song window

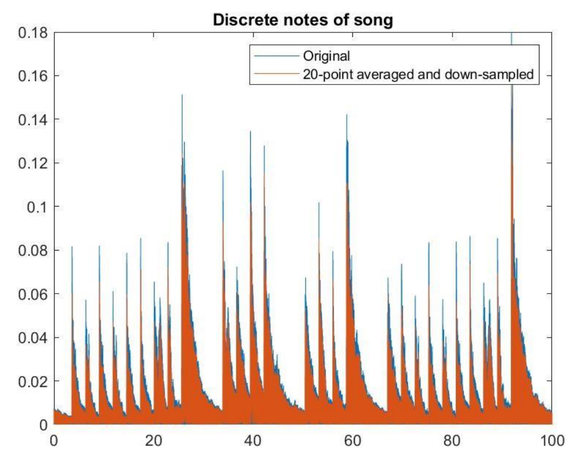


Fig. 4. Averaged and Downsampled Respresentation of Audio Signal

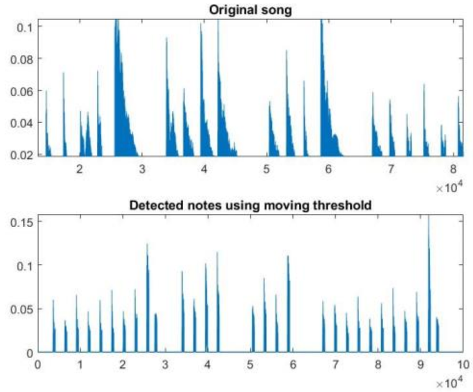


Fig. 5. Original Signal and Signal after Thresholding

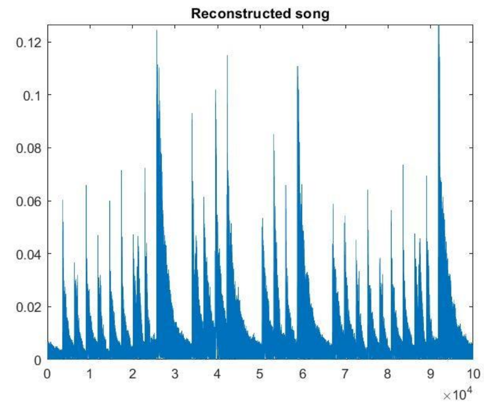
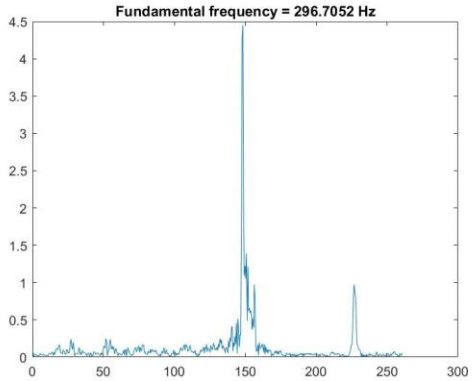


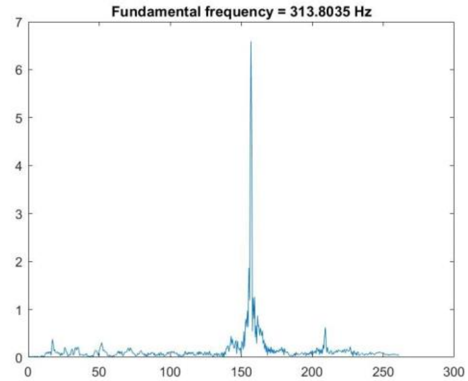
Fig. 6. Reconstructed Audio Signal

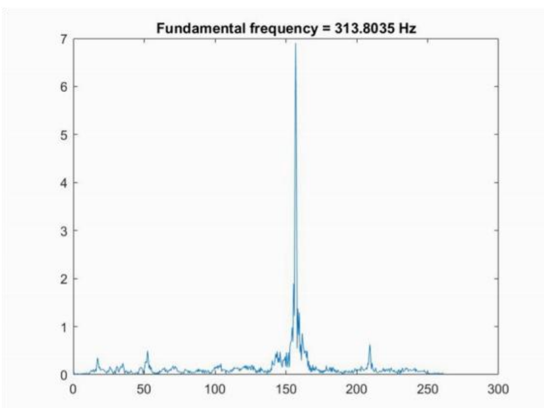
```
letter =

  29×1 string array

    "Eb4"
    "D4"
    "Eb4"
    "D4"
    "Eb4"
    "Bb3"
    "Db4"
    "B3"
    "Ab3"
    "Ab2"
    "B2"
    "Eb3"
    "Ab3"
    "Bb3"
    "Eb3"
    "G3"
    "Bb3"
    "B3"
    "Eb3"
    "Eb4"
    "D4"
    "Eb4"
    "D4"
    "Eb4"
    "Bb3"
    "Db4"
    "B3"
    "Ab3"
    "Ab2"
```

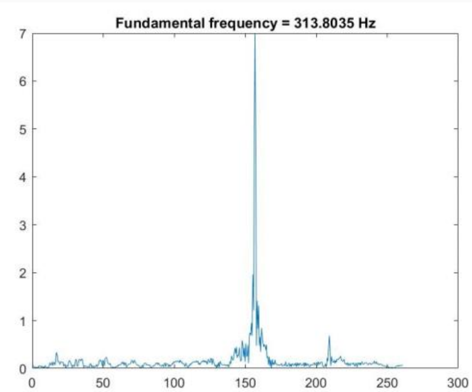
Fig.7 Output of Notes

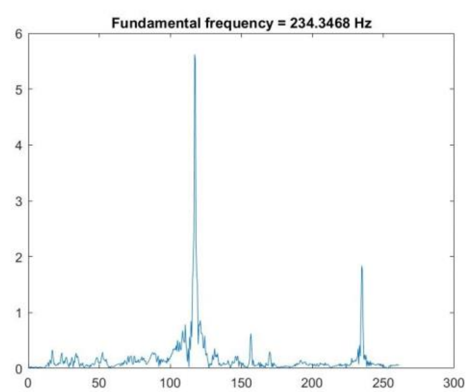
Below are the individual frequencies present in the song

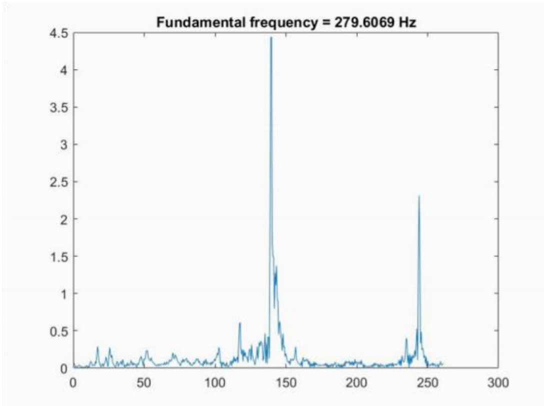


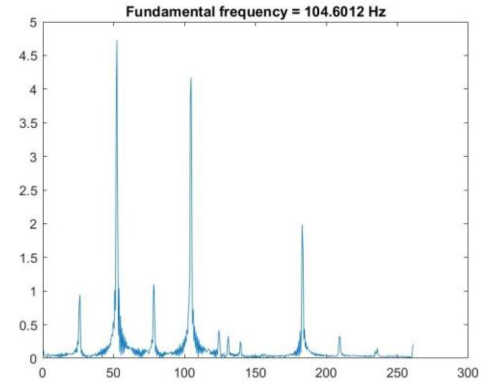


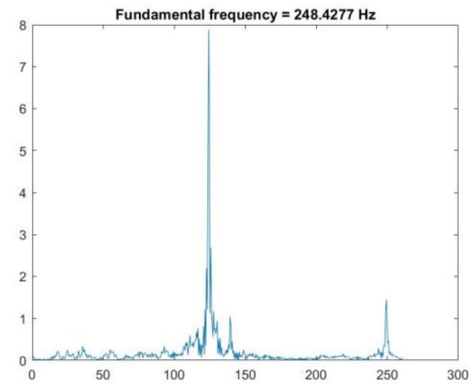


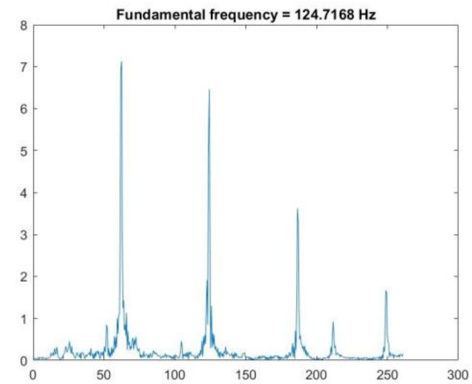


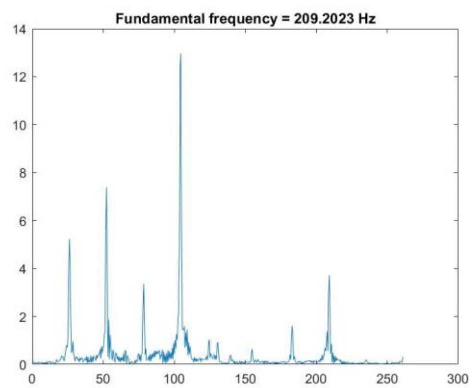


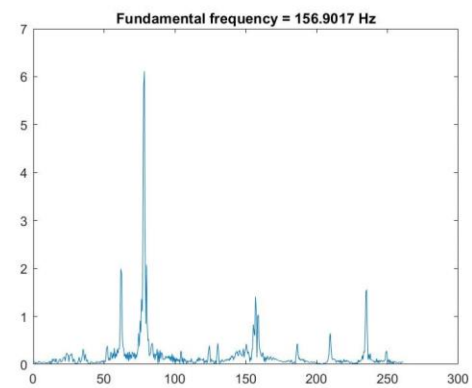

Fundamental frequency = 279.6069 Hz

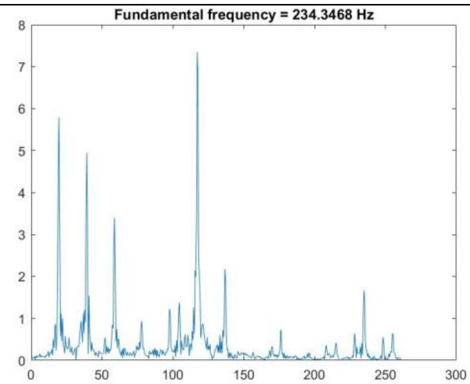

Fundamental frequency = 104.6012 Hz

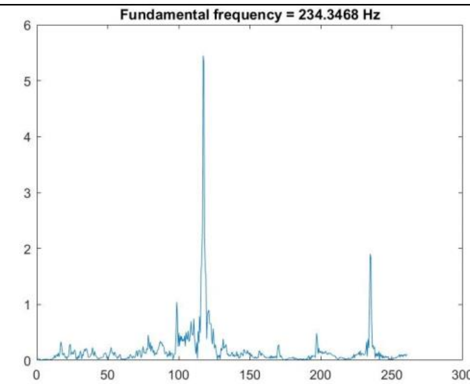

Fundamental frequency = 248.4277 Hz

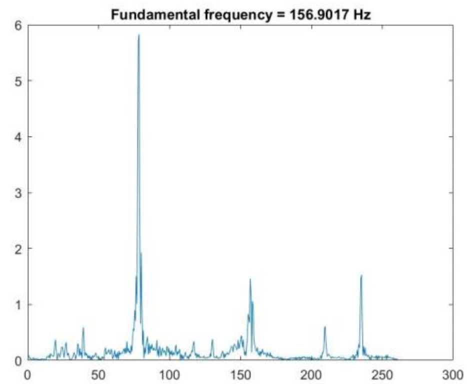

Fundamental frequency = 124.7168 Hz

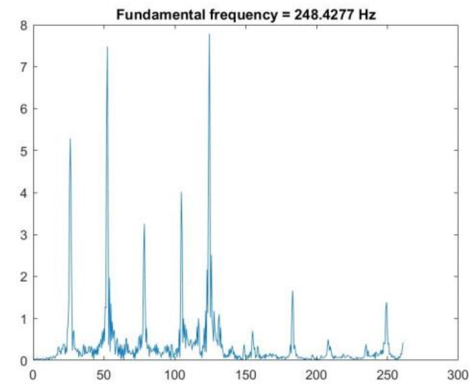

Fundamental frequency = 209.2023 Hz

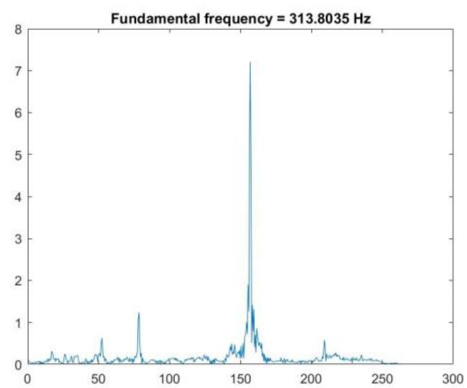

Fundamental frequency = 156.9017 Hz

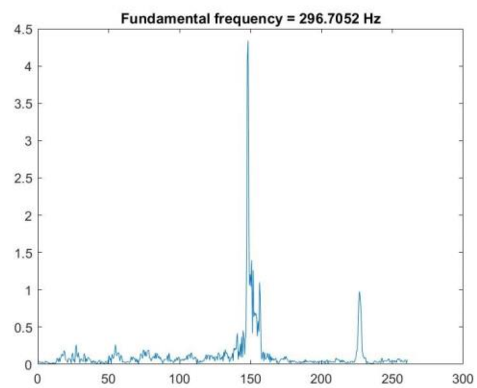

Fundamental frequency = 234.3468 Hz

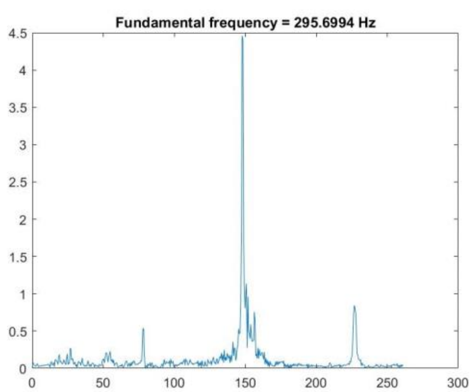

Fundamental frequency = 234.3468 Hz

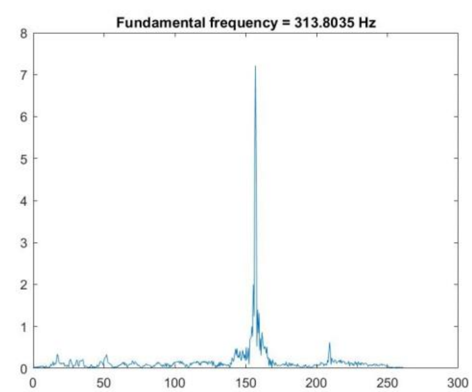

Fundamental frequency = 156.9017 Hz

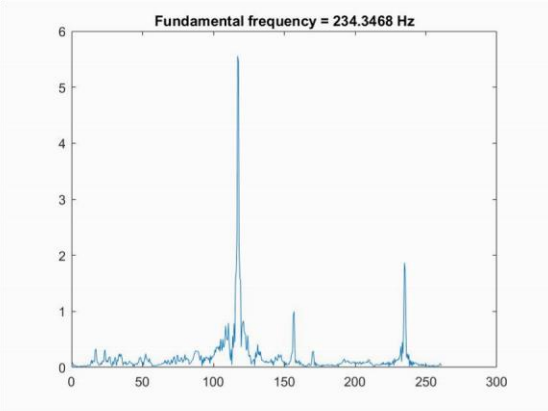

Fundamental frequency = 248.4277 Hz

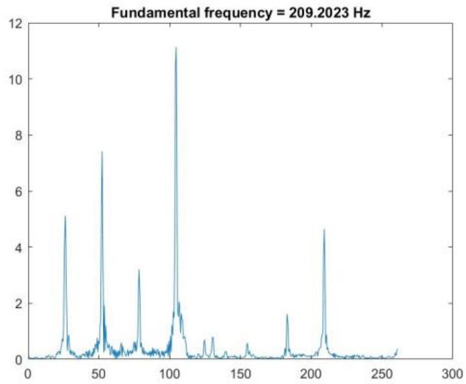

Fundamental frequency = 313.8035 Hz

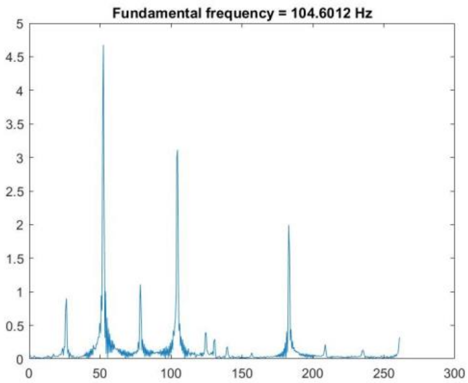

Fundamental frequency = 296.7052 Hz


Fundamental frequency = 295.6994 Hz


Fundamental frequency = 313.8035 Hz







Results

Happy Birthday, Jingle Bell, Twinkle-Twinkle, Fur Elise, and other songs were used in the experiment. Fur Elise's findings, which are the audio sample's notes, are presented below. Therefore, the notes on the piano are detected for the Audio Signal based on Fur Elise is, 29 notes were detected by the recognition algorithm. The fundamental frequencies are plotted us. The reconstructed audio signal reconstructed from the detected audio signals has following waveform.

Conclusion

This study proposes a method for automatically recognising music notes. The Fast Fourier Transform is used to convert time domain signals into frequency domain signals with continuous proportional resolution. Template The note is then detected and identified using matching and other processing (s) Electronically synthesised and actual piano pieces are used to test the algorithm. Perfect recognition is achieved for the synthesised pieces examined. High precision is also obtained for the genuine piano works. If the notes are played exactly according to the score, more precision can be attained.