AUTOMATIC IDENTIFICATION OF FREQUENCIES OF MUSICAL NOTES IN POLYPHONICS WAVE FILES

Carlos Alberto Lopes da Silva
Federal University of Uberlândia
Brazil
carlossilva@quantica.com.br

Gilberto Arantes Carrijo
Federal University of Uberlândia
Brazil
gilberto@ufu.br

Luciano Vieira Lima
Federal University of Uberlândia
Brazil
rhodan@quantica.com.br

Antônio Claudio P. Veiga
Federal University of Uberlândia
Brazil
acpveiga@ufu.br

Abstract

This paper deal with the problems that can be found when someone wishes automatically identify frequencies of musical notes recorded in WAVE format. The same approaches the problem of the analog to digital conversion and of the sample rating of the files in this format when analyzed through Fast Fourier transform. This transform generates a significant amount of lateral lobes around the existent frequencies in these files, taking the professional to identify musical notes not existent in the original files. Like this, this paper presents optimum values of windows to minimize such problem. An appropriate window is not just an efficient and definitive solution to this project. This way a forecast method, refinement and identification is presented here to identify musical notes in files in the WAVE format .

1 - Introduction

The man is always looking for automatize tiresome and difficult accomplishment tasks. So, he has been looking for mathematical tools and heuristics that qualify the automatic processes to accomplish these tasks that he judges simple or that he can't do.

The fact of the man examines the sky and to recognize and to differentiate the planets and stars are difficult tasks of being automated. One of the existent reasons is due to the limitations imposed by the acquisition systems and digital conversion of data and of the existent mathematical tools to manipulate such data. While the eye and the human ear process continuous signals, a computer depends of given digital data that are far away to be considered satisfactory for an efficient analysis.

One of these problems is the transformation of a sound file in the WAVE format for the MIDI format. The problem is in the recognition of the frequency and duration of musical notes in a file WAVE and code in a file in the MIDI format SMF 0 or 1.

2 - The great problem

To discover new stars in the sky has everything in common with discovering musical notes in certain music. In both cases we have continuous signals, in other words, the light of the stars and the sound of music, where the main problem is that it's not known what it looks for. For this case the existent frequencies of the stars are not known sought in the sky as well as the frequencies of the musical notes of the sound signal. The task gets complicated when we tried to apply FFT in the signals. This tool, when applied to signals whose frequencies are not known, they produce frequencies non-existent in the original signals as a result of the analysis. Mello [1] in your paper “Estimation of periods from unequally spaced observations” for The Astronomic Journal - 1981, got to idealize the first mathematical tool to solve the problem of identification of frequencies of stars ignored in a digital map of any section of the sky. Silva et all [2] it sketches in your paper for Diderot 1999 a migration of such technology for a possible solution, also, of the problem proposed in this paper.

3 – The problem for recognition of the musical notes

To be possible to analyze a sound file correctly it is initially necessary to divide it in pictures (windows) that contain 2^n points in a signal period, where n is the total number of points of the picture. Such division is necessary to use Fast Fourier Transform (FFT).

When run FFT with a window of fixed size in a file containing a musical note, the detection of this note is restricted to a maximum size of the window. This is characterized due to the fact that the analog to digital conversion introduces mistakes in the file (aliasing). When analyzing a file WAVE containing only a musical note (C4) we can observe the following facts: when we apply FFT with a picture size with more than 8192 samples, the musical note C4 is found and mistakes are not introduced (presence of other notes). When the size of the picture is reduced (4096–2048–1024, etc) notes that don't belong to the file starts to appear. This fact is possible of being visualized through the table 1. With this we ended that, starting from a size of window of 4096 bytes, not more it is possible to identify a musical note accurately. It can be ended that, for this case, the ideal size of window should be 8192 bytes, too big for a detailed analysis of all the notes present in the file.

Original note	**Number of Pictures for Sample**	**Picture**	**Found notes**		**WINDOW (Bytes)**
			Note	**Intensity (dB)**	
C4	01	01	C4	114,0	16384
C4	02	01	C4	113,0	8192
		02	C4	112,0	8192
C4	04	01	C4 Db4	113,0 109,0	4096
		02	C4 Cb4	114,0 111,0	4096
		03	C4 Db4	112,0 109,0	4096
		04	C4 Db4	111,0 108,0	4096
C4	08	01	Db4 D4	110,0 107,0	2048
		02	Db4 D4	110,0 106,0	2048
		03	Db4 D4	111,0 108,0	2048
		04	Db4 D4	109,0 105,0	2048
		05	Db4 D4	109,0 105	2048
		06	Db4 D4	108,0 105,0	2048
		07	Db4 D4	108,0 105,0	2048
		08	Db4 D4	108,0 105,0	2048

Table 1–musical Notes found for the file C4 being varied the size of the sampling window.

4 - Before and after processing filters

To analyze and to identify musical notes in an audio digital file has a great advantage when compared to analyze and to identify stars in a digital picture of the sky. In the case of the stars they can assume any inside frequency of a strip of the visible spectrum. In the case of the music using the very temperate scale this is not possible. A discrete number of possible musical notes exist inside of the audible sound range so that this range is not continuous, in other words, between a frequency and its next exist a relation of 1,059463. Another difference is that the stars can be located anywhere and in any arrangement. This doesn't happen in the harmonic traditional music with tuned musical instruments. In this case, the musical notes should obey a harmony that limits the distance among them and, also, the form with that can group. Like this, several restrictions exist imposed by the musical domain, the ones which, if used in the algorithm of identification of the frequencies of musical notes, they will significantly be able to improve the indexes of successes of the researched notes. Some procedures should be accomplished before the beginning of the analysis as filters of the analyzing system. Other procedures will be used after having accomplished the analyses and identified the possible existent musical notes. Due to the fact of the algorithm can identify, in special circumstances, nonexistent notes, the post processing filters will be very useful to avoid that erroneous identifications happen.

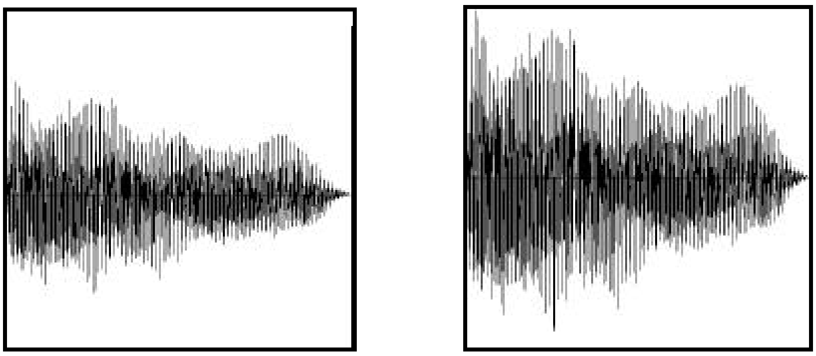
5 - The algorithm of detection of musical notes

The algorithm is subdivided in 3 parts.

1- Normalization of the signal.
2- Forecast of the possible existent musical notes in the files WAVE.
3- Refinement and identification of the existent musical notes.

5.1 - Normalization

The normalization will be made by the value of pick of the signal for 0db. As the proposed algorithm is not concerned in identifying, initially, the volume of the existent musical notes in the files, this artifice guarantees that every existent real note in the file will possess a close volume of zero db.



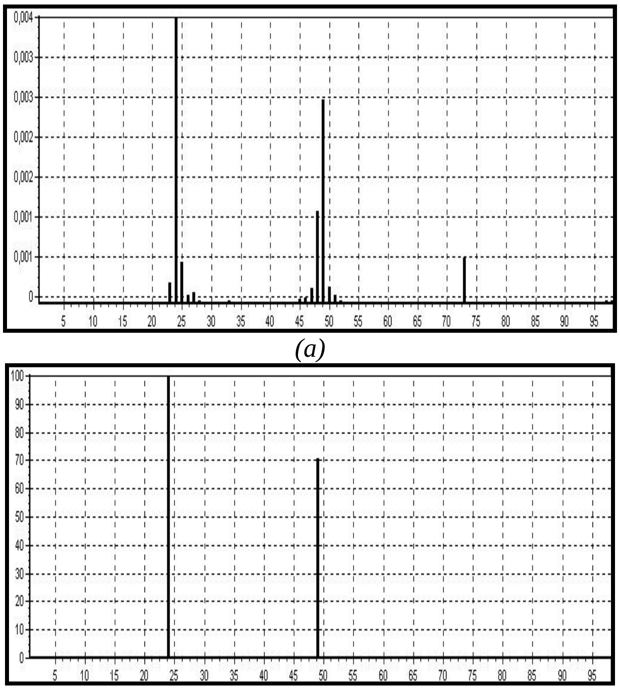
(a) (b)

Illustration 2 – Form of wave of a note C4. (a) Without normalization (b) With normalization

5.2 - Forecast

In this stage we will try to foresee which are the possible existent notes in the analyzed file. This forecast is made through the analysis of Fourier. With this analysis they are eliminated the picks that possess an inferior volume at the defined minimum level for the user. We can divide this stage in two other:

a) Filtered Identified frequencies due to analog to digital conversion mistake and for ignorance of the period that possess levels in decibel below to the expected ones in the file, since the same is normalized in 0 db.



(a)

(b)

Illustration 3–Spectrum of Fourier of a note C4. (a) Without Filter (b) With Filter and normalized from 0 to 100

b) Filtered Identified frequencies through the filter of notes of the very temperate scale. In this paper we will assume that the instruments are tuned with the note A 440Hz. For the example of the Illustration 3b, the existent pick in the position 24 (a) is equal a frequency of 258,4 Hz. Being consulted a table of tuned musical notes is possible to observe that it approaches of the frequency of 261,63 Hz, that is a C4. The existent pick in the position 49 (a) is equal a frequency of 527,6 Hz. Being consulted the same table of tuned notes is possible to observe that it approaches of the frequency 523,25 Hz, that is a C5. Possibly this second notice found it should be a harmonica of the fundamental.

The expected frequency is obtained through the following formula:

$$F = \frac{(i * Fs)}{N}, \qquad (1)$$

Where i is the reference index in the vector of FFT, Fs is the sampling frequency and N is the number of points of FFT. For the mentioned example the sampling frequency is 44100 Hz and the size of the sample is 4096.

5.3 - Refinement and Identification

Made the forecast of the possible notes, the spectrum obtained is already much cleaner, just containing possible musical notes. It remains to know if the foreseen notes really belong to the analyzed file. The main reason of the identification of nonexistent notes in the original acoustic sound is that in the analysis of Fourier of the sample signal it is being applied in a window whose number of points doesn't correspond a period of the signal, happening this way an aliasing mistake. Another factor is that it is not had 2^n points in the period analyzed by the algorithm of FFT and, this way, it is not possible the exact identification of the fundamental frequency. This module, refinement and identification, takes charge of solving this problem.

Taking the foreseen notes, it owes you are processed them to verify your pertinence to the file. That is done being determined the number of necessary points for a period of signal. This number is calculated through the equation:

$$NP1 = \frac{SR1}{F1}, \qquad (2)$$

Where SR1 is the frequency of sampling of the signal, F1 it is the frequency of the signal and NP1 it is the number of necessary points for this signal. With this value we can calculate a value NP2 that contains 2n points the closest possible of NP1. The ideal case is when NP2 is smaller or equal NP1. This way the aliasing mistake is avoided, being eliminated discharges frequencies that could be introduced. Of ownership of the value NP2 and using the equation 2 a new sampling frequency SR2 is calculated for the signal so that it can be made a new sampling of the signal through the interpolating use. The illustration 4 illustrates this process.

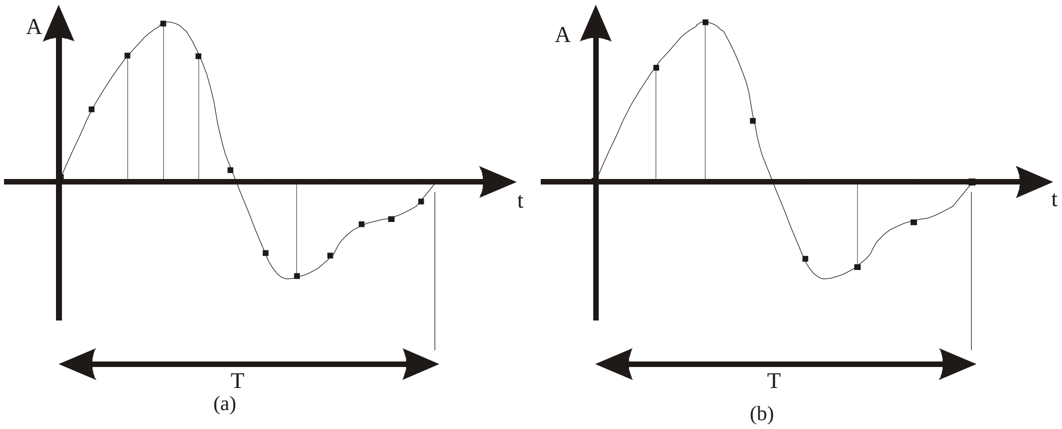


Illustration 4– resample of an audio signal. (a) Period containing 12 samples (b) Period containing 8 samples.

Illustration 5 displays the process of signal interpolation:

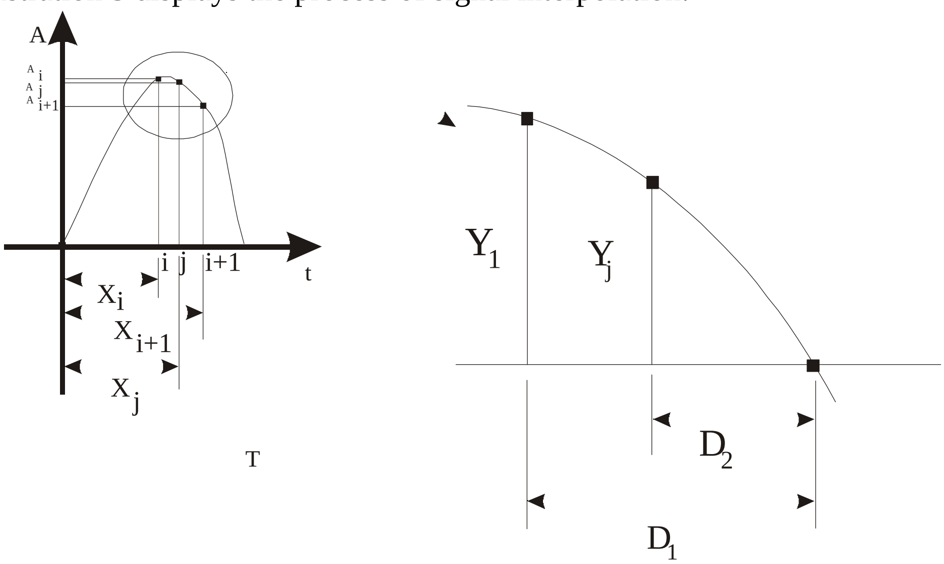


Illustration 5–Interpolating with different sample rates

$$Y_1 = A_1 - A_{1+1} \qquad (3)$$

$$X_i = \frac{i}{FS1} \text{ e } X_{i+1} = \frac{i+1}{FS1} \qquad (4)$$

$$X_j = \frac{j}{FS2} \quad (5)$$

$$D_1 = \frac{1}{FS1} \text{ e } D_2 = X_{i+1} - X_j = \frac{i+1}{FS1} - \frac{j}{FS2} \quad (6)$$

$$Y_j = (A_i - A_{i+1}) * FS1 * \left(\frac{i+1}{FS1} - \frac{j}{FS2} \right) \quad (7)$$

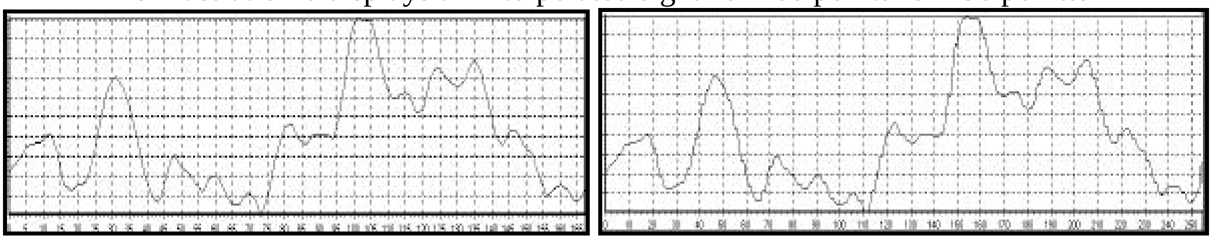
$$A_j = Y_j + A_{i+1}, \quad (9)$$

to $1 \le i \le FS1$, $1 \le j \le FS2$ e $|A_{\succ}| \succ |A_{i+1}|$. To $|A_{\succ}| \prec |A_{i+1}|$, we have:

$$Y_j = (A_{i+1} - A_i) * FS1 * \left(\frac{j}{FS2} - \frac{i}{FS1} \right) \quad (10)$$

$$A_j = Y_j + A_i. \quad (11)$$

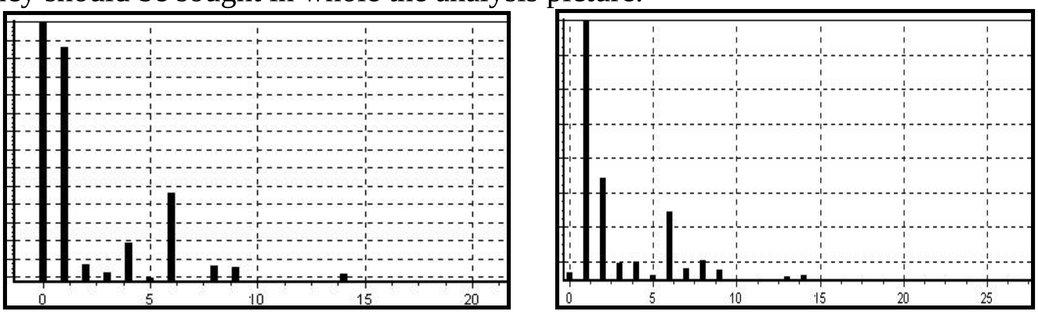
The illustration 6 displays an interpolated signal of 168 points for 256 points.



(a) *(b)*

Illustration 6–(the) 168 points of a signal C4, corresponding to a period (b) Interpolating of 168 points for 128 points of the same signal C4, corresponding to a period

With this interpolating, when applying the FFT in this signal, if the dear frequency really exist in the file, it will appear a peak in the fundamental of the transformed, what will confirm your existence. If we try to find non-existent note, the pick in the fundamental of the transformed it will be zero or very small, as display the illustration 7. The existence of this frequency should be sought in whole the analysis picture.



(a) *(b)*

Illustration 7–(a) Spectrum of Fourrier in the search of a note C4 (b) Spectrum of Fourrier in the search of a note D4

Observe that in the illustration 7a we are trying to find a musical note. We previously know that this is the note C4, what was proven for the algorithm, in other words, we obtained a maximum peak in the fundamental (index 0). When we sought a musical note that it doesn't exist in the file, its fundamental frequency won't appear in the spectrum, what is showed through the illustration 7(b).

With this process it is possible to determine the existence of one or several musical notes exactly in an interval of time. It is enough now to determine the location and exact duration of these musical notes.

In first place, it should be made a sweeping in whole the picture, being looked for the existence of the notes foreseen music, as well as your beginning and duration. For this it is taken as reference the smallest possible resolution for analysis being made a sweeping in whole the picture and applying again a FFT, trying to identify only the existence of the frequencies foreseen previously. In case of frequency or frequencies appear in at least two thirds of the minimum resolution, it is considered as being a musical note. This way we can be known if the note it really exists and where will be located. The illustration 8 shows the exposed above:

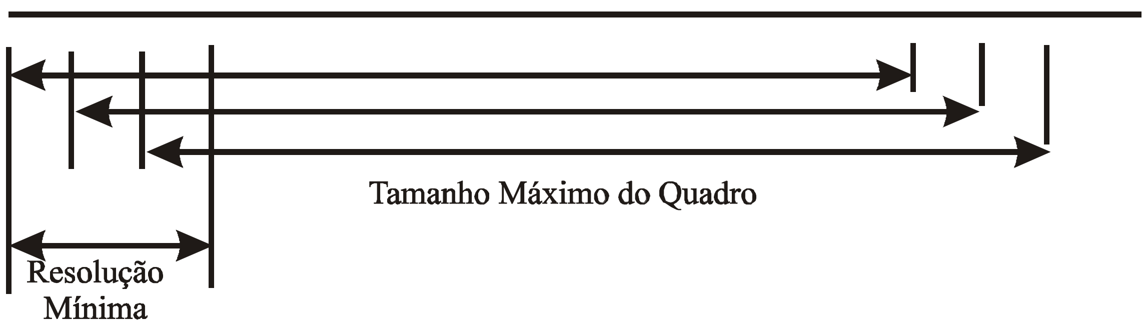


Illustration 8–sweeping Form in the picture of samples

6 - Future works

This project is part of main system responsible for identify musical notes in polyphonic files in WAVE format project with subsequent conversion for the Standard MIDI File (SMF) format 0 and 1. For this conversion, the largest barrier was in being determined when a note begins and when it ends, to know a new note is beginning or if she is already harmonic of other existent. In future goods the algorithms of temporary and dynamic identification of the notes will be presented in the files in the format WAVE and the process of conversion of the same ones in SMF.

7 – Concluding Remarks

The proposed algorithm is quite faithful in the detection of the frequencies of the musical notes registered in a Wave format file. The artifice of interpolate and reorganize the points in the interval of sampling of the note foreseen it really serves as a good pre-processing filter avoiding the use of the referred post processing filters. Such filters will be, without any doubt, important in the case of polyphonic complex music executed for more than a musical instrument simultaneously.

8- References

[1]–MELLO, S. FERRAZ. “Estimation of Periods from Unequally Spaced Observations”, The Astronomical Journal–Volume 86, Number 4–April, 19981.

[2]–HE/SHE WHISTLES, CARLOS A. L. OF THE, PASCHOARELLI, ANTÔNIO CLÁUDIO, CARRIJO, GILBERTO ARANTES, IANO, YUZO, “Use of Fast Wavelet And Fourier Transforms In Music Recognition”–DIDEROT FORUM on Mathematics and Music - LISBON - PARIS - VIENNA, 3rd and 4th December - 1999.