IDENTIFICATION OF RHYTHM AND SOUND IN POLYPHONIC PIANO RECORDINGS

PACS REFERENCE: 43.75.Mn

Ana María Barbancho Pérez, Antonio Jurado Navas, Isabel Barbancho Pérez
Departamento de Ingeniería de Comunicaciones. Universidad de Málaga
E.T.S. Ingeniería de Telecomunicación. Campus Universitario de Teatinos s/n
29071 Málaga
Spain
Tel: +34 952 13 41 62
Fax: +34 952 13 20 27
E-mail: abp@ic.uma.es

ABSTRACT
A system that is capable of identifying all the elements in a polyphonic piano recording which are necessary to write out the score, using a record of the musical piece, is presented in this contribution. A piano piece is processed and analysed using wavelet filters and the discrete Fourier transform, the correlation properties among the different harmonics of the sounds are also used to identify its rhythm and sound. By this means, the system designed identifies the different notes that were played, their attack instant, span and intensity. An analysis if these results allows to estimate the key signature, tonality, tempo and time signature of the piece.

RESUMEN
En esta comunicación se presenta un sistema que es capaz de proporcionar todos los elementos necesarios para escribir la partitura de una obra polifónica de piano, a partir de una grabación de la misma. Para llevar a cabo la identificación del ritmo y sonido en la obra pianística, la señal musical es procesada y analizada utilizando filtrado wavelet, la transformada discreta de Fourier y las propiedades de correlación existentes entre los distintos armónicos de los sonidos. De esta forma, el sistema diseñado determina las distintas notas ejecutadas, su duración, su intensidad y el instante de ejecución. Analizando estos resultados, se calcula la armadura, la tonalidad, la modalidad, el tempo y el compás de la obra musical.

1. INTRODUCTION

A score or a performance allows a musical piece to be identified. In the score, the different sound parameters are described, through musical notation [1]. The musical symbols that make up the score are clefs, notes, key signature, time signature and so on. On the other hand, the musical piece performance, that is, the sound that is enjoyed, can be described by the set of frequencies that appear along the time with certain intensity.

Nowadays, there exist several computational systems capable of playing a musical piece from its score. However, the opposite process to that, the extraction of the score from the musical signal by computational transcription systems, is not satisfactory solved yet [2]. In this paper, a system that obtains all the musical elements to transcript the score, from a piano polyphonic record, is presented. This system has been designed to be independent of the performer and the piano used in the recording and does not require a training process at all.

This paper is structured as follows. In section 2, the system design is presented and behaviour and relationship among the different stages is studied. Afterwards, some results obtained by the system in different piano pieces are show in section 3. Finally, some conclusions are drawn in section 4.

2. SYSTEM DESCRIPTION

The polyphonic identification system is composed of three main blocks, as shown figure 1. In the first stage, the piece is divided into a number of time intervals. To perform this division, the system locates the time instant when one or several simultaneous notes are played. In other words, each attack instant is located and its energy and its duration are obtained. Then, the second stage identifies the notes played in each time interval, thus, the system obtains the notes played along the piece, together with their duration, intensity and the time instant when they were played. Using all this information, the final stage calculates the musical parameters of the piece: key signature, tonality, tempo and time signature.

The following sections describe in detail the techniques used in the different blocks of the system.

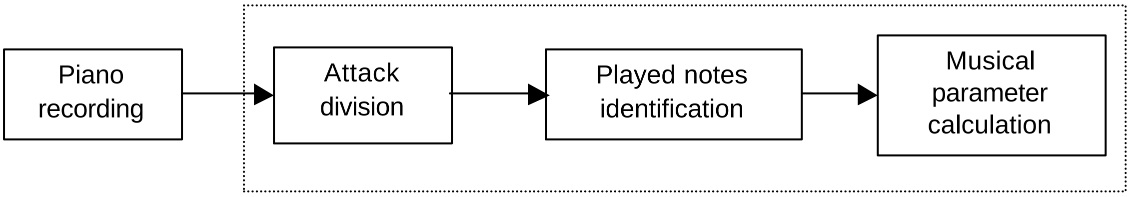

Figure 1. Polyphonic identification system.

2.1. Attack division

A sudden energy increase happens when a piano note is played. This event is called an attack [3]. This feature is used to divide the piece into time slots. Each slot or interval starts when a single note, or several simultaneous notes are played. Each part of the piece obtained is called attack slot or attack interval. Figure 2 presents an example of attack division, the sudden energy increase that appears after each touch is clearly observed. Attack detection is done using sliding windows. In each window, a test is performed to detect an energy gap [4]. The length and the intensity of each attack are stored since they will be used in the following stages.

2.2. Played notes identification

Played notes identification is performed in each of the attack slots detected in the first stage. The process is based on the identification of the fundamental frequencies of the notes played. To that end, pedal detection is performed in each slot, afterwards, the slot is analysed in a filter bank. The filter bank finds the frequency band in which the played notes are located, then, the frequencies that exceed a certain energy threshold are found. The thresholds used in the detection depend on the results of the pedal detection stage. Using information from the previous attack slot and after a number of tests, the fundamental frequencies are selected among the frequencies that passed the energy test. Next, the parts that make up that stage (see figure 3 for a scheme) will be described.

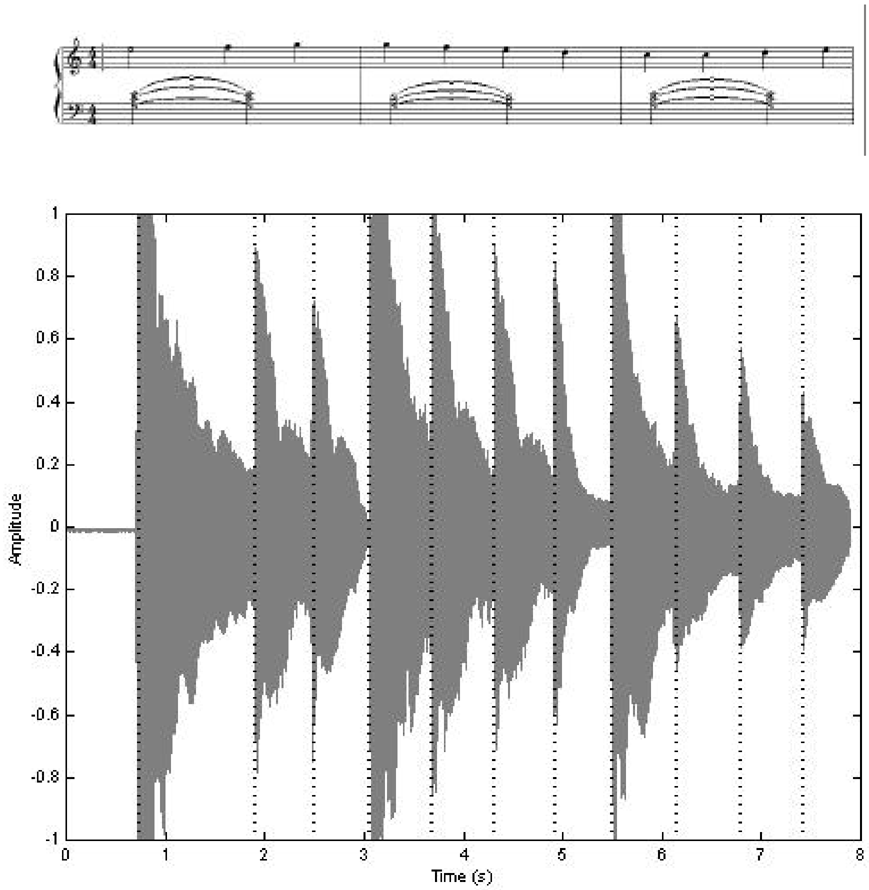

Figure 2. Attack division.

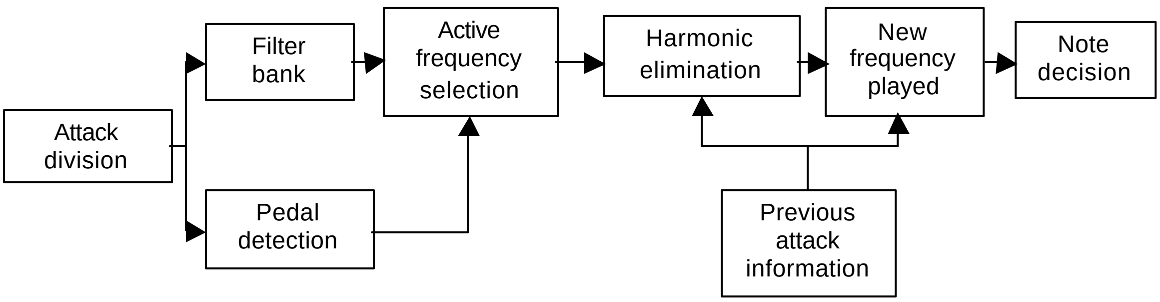

Figure 3. Scheme for the identification of played notes.

a) Filter bank processing. Each attack slot is analysed by a five-stage tree-structured filter bank [5]. We look for the subband that contains 90% of the energy. The bandwidth of the subband selected depends on the number and the types of the notes that were played.

b) Pedal detection. When the piano pedal is pressed, the noise floor is increased and there appear additional frequencies in the spectrum that could be considered played notes. Pedal detection is done averaging the module of the spectrum of the signal.

c) Frequency selection. The frequencies that exceed a certain threshold in the module of the spectrum are considered the active frequencies (figure 4). Three different thresholds are considered: one for low frequencies (0-200Hz), another one for medium frequencies (200-800Hz) and the last one for high frequencies (>800 Hz). A detailed analysis of a number of pieces has been performed to define the thresholds. Note that threshold selection also depends on the use of the piano pedal. An example of frequency selection is presented in figure 4. Several thresholds are defined to select the frequencies using the module of the spectrum. In this case, the frequencies selected correspond to C3, D#3, E3 and E4.

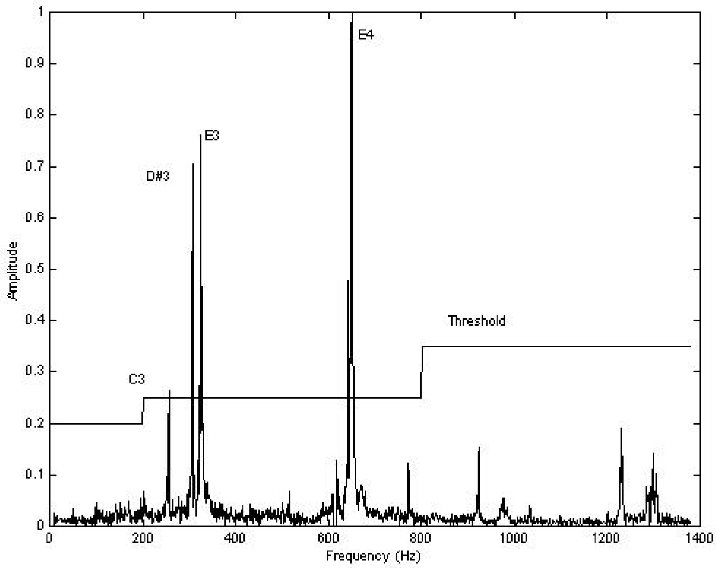

Figure 4. Frequency selection.

d) Elimination of harmonic frequencies. Before the process of note detection, it is necessary to make sure that the frequencies to analyse correspond solely to the fundamental frequencies of the played notes and not to their harmonic frequencies. So, a frequency is considered to be a fundamental frequency if it passes all of the following elimination tests:

- Amplitude elimination test. The amplitude of each frequency is compared with a threshold that is proportional to the amplitude of the frequency that could be its fundamental frequency. If the threshold is higher, then that frequency is considered a harmonic frequency and it is removed.

- Bandwidth test. It's has been empirically checked that the piano notes are such that the full width half maximum (FWHM) bandwidth of their harmonic frequencies is the same as the FWHM bandwidth of the fundamental one. So, the frequencies that satisfy this condition are also eliminated.

- Inter-attack energy test. The harmonic frequencies obtained in an attack interval which do not pass an energy test relative to their energy in the previous attack interval are considered harmonic frequencies and then they are removed.

e) Detection of new frequency played. This part decides if the identified frequency is newly played in this attack or it is slurred to the previous ones. This information is extracted from the correlation between a sinusoidal reference signal of the frequency under study and the previous and actual attack interval signal [6]. If the correlation has no discontinuities (figure 5a), the notes was played in a previous attack, otherwise, it's a new note (figure 5b).

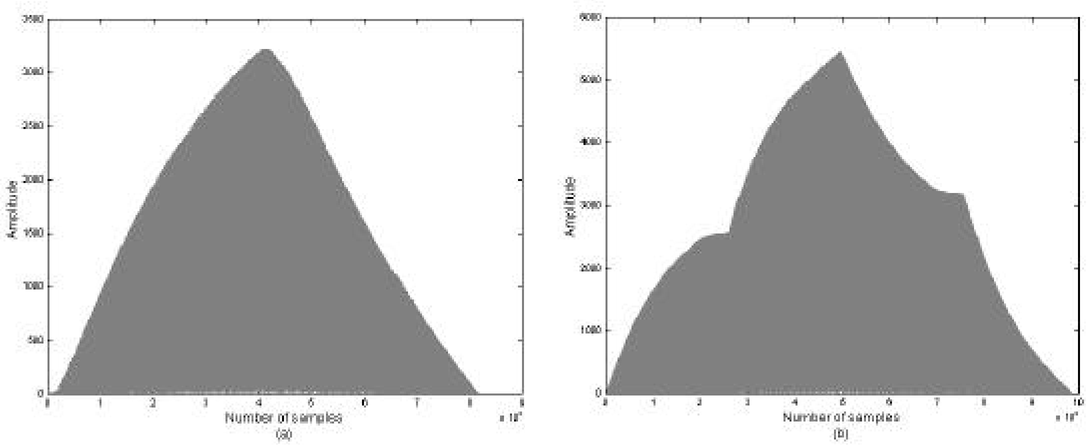

Figure 5. Correlation plots. a) Note played in the previous attack. b) New note.

f) Note detection. Finally, the decision about the notes and the octaves that correspond to the selected frequencies is made.

2.3. Musical parameter calculation

In the previous sections, the algorithms to detect the played notes and their length and intensity have been presented. Now, the key signature, tonality, tempo and time signature of the considered piano piece will be calculated.

2.3.1. Key signature and tonality

Key signature calculation is based on the distribution of the accidental notes in the piece. An algorithm that minimises the number of accidental notes is employed. After deciding the key signature, the tonality must be calculated. To distinguish between major and minor keys, the number of times that the leading tone is separated a halftone from the tonic, assuming minor key, is counted. If in more that 50% of the cases the leading tone is separated a halftone from the tonic, the system selects minor key, otherwise, the system decides major key.

2.3.2. Tempo estimation

For tempo estimation, the time selected is the most frequent length of the attack slots along the piano piece. The note value assigned is based on the attack length (table 1). These temporal values have been obtained empirically after studding several musical pieces.

Attack length	>0.75s	>0.4s y <0.75s	<0.4s
Note values	𝅗𝅥	♩	♪

Table 1. Relationship between note value and pulse length.

2.3.3. Time signature calculation

Time signature calculation is performed in two stages. In the first stage, the most frequent length of the attack slots defines the bottom number of the time signature. During the second stage, the top number of the time signature is obtained testing all the possible time signatures, then, the average energy of the first attack of each measure is calculated. The system selects the time signature that maximises this parameter.

3. RESULTS

In this section we present some of the results obtained by the system described in this contribution (table 2).

Musical piece record	*Attack detection*	*N*	*FN*	*FP*	*Score*	*Time signature*	*Key signature*
Beethoven. Sonata nº 9. 2nd movement	98.75 %	267	13	40	83.48%	3/4	E minor (right)
Scott Joplin. The Sting	100 %	102	10	12	82.26%	8/8 (right 4/4)	C major (right)
Beethoven. Minuet in G major	93.3 %	75	8	10	80.65%	3/4	C major (G mayor)
Ana Magdalena Bach. Notebook 20 pieces. Minuet nº 20	97.14 %	67	5	7	84.81%	3/4	F major (right)

Table 2. Systems results

In table 2:

- N is the number of notes identified correctly.
- FN stands for the number of played notes that were not detected by the system.
- FP is the number of notes reported by the system that were not played.
- Score is calculated as:

$$PA = \frac{N}{N + FN + FP}$$

If we compare this result (table 2) with the results obtained by the system described in [7], in which the score is 70%, we can appreciate the advantage of our system. Besides, notice that the system in [7] must be trained and it detects neither the key signature nor the time signature. Concerning the error found in the estimation of the time signature in “The Sting”, it is easily justified since the predominant note in this piece is the quaver. With regard to the error in the detection of the tonality in Minuet in G major, it must be pointed out that the difference between the selected key signature and the real one is only an accidental. Furthermore, some modulations towards the tonality detected occur in this piece, so the systems detects that the piece is longer in C major than in G major.

4. CONCLUSIONS

A system that is capable of determining all the elements that are necessary to write out the score of a polyphonic piano piece form a record is presented. The system identifies the notes, their length, their intensity and the time instant when the notes were played. Additionally, the system finds the key signature, the tonality, the tempo and the time signature of the piece under analysis. Wavelet filters, the discrete Fourier transform and the correlation properties among the different harmonic frequencies of the notes are the tools used by the system designed.

5. BIBLIOGRAPHICAL REFERENCES

[1] D. de Pedro, “Teoría de la Música. Vol 1 y 2 ”, Real Musical 1992.

[2] E. Franov, “¿Hacer música?. ¡Con MIDI es más fácil!”, Magazine Casinada. http://usuarios.iponet.es/casinada/18midi01.htm.

[3] K. Jensen, “Envelope model of isolated musical sounds”, Proceedings of the second COST-G6 Workshop on Digital Audio Effects, Trondheim, December 1999.

[4] S. Dixon, “Learning to Detect Onsets of Acoustic Piano Tones”, MOSART Workshop on Current Research Directions in Computer Music, Barcelona.

[5] M. Bobrek, D.B. Koch, “Music signal segmentation using tree-structured filter banks”, Journal on Audio Engineer Society, Vol. 46, No. 5, May 1998, pp. 412-427.

[6] T. Schroeter, S. Doraisamy, S.M. Rüger, “From raw polyphonic audio to locating recurring themes”. Proceedings of the International Symposium on Music Information Retrieval, Plymouth, Octubre 2000.

[7] S. Dixon, “On the computer recognition of solo piano music”. Mikropolyphonie, Vol. 6, 2000.