ARCHIVES OF ACOUSTICS
32, 3, 617–629 (2007)

APPLYING COMPUTATIONAL INTELLIGENCE TO MUSICAL ACOUSTICS

Bożena KOSTEK

Gdańsk University of Technology
Multimedia Systems Department
Narutowicza 11/12, 80-952 Gdańsk, Poland
e-mail: bozenka@ssound.eti.pg.gda.pl

(*received May 14, 2007; accepted July 4, 2007*)

The aim of this paper is to review some selected computational techniques that find application in acoustics and in particular to sound engineering. The presented research studies involved using artificial neural networks, rough set method, fuzzy logic, genetic algorithms and other soft computing techniques. The investigated problems are related to classification of musical instrument sounds, musical phrases recognition, intelligent music processing, computer control of classical pipe organ instruments, and quality assessment.

Keywords: musical acoustics, musical informatics, music information retrieval, soft computing, classical pipe organ.

1. Introduction

The experiments carried out in the Multimedia Systems Department (former Sound and Vision Engineering Dept.) of the Gdańsk University of Technology in the field of acoustics involved *computational intelligence* and *soft computing* algorithms for many years [6, 25]. This discipline, which goes beyond traditional concepts of artificial intelligence, experiences rapid development now. It comprises algorithms of artificial neural networks, fuzzy logic, rough sets, genetic algorithms, decision trees, classifiers based on the nearest neighborhood method, and hybrid methods involving two or more of the above listed approaches. The mentioned methods also take cognitive approach to human perception into account [7].

The paper recalls some chosen problems that required solutions employing methods of computational intelligence. For convenience, this work also shortly reviews notions or theoretical basis of selected algorithms from the field of artificial intelligence. Among others, the implementation of intelligent methods to control classical organs extended the set of their applications to acoustics. The last mentioned topic is of particular interests to both musical acoustics field and soft computing methods. It should be mentioned that musical acoustics supported by the artificial intelligence domain evolved into a new domain, called musical informatics.

The starting point of any study should be a thorough research in the area of interests resulting in a collection of some key papers in the domain. Some articles and papers written by Professor Andrzej RAKOWSKI [13–17] became inspirational guidance to the author of this paper, especially those related to study on musical timbre and organ sounds in particular. As also happened that he was the Reviewer of both author's Ph.D thesis and her habilitation dissertation. Professor Rakowski was kind enough to appreciate the novelty of these topics even if at that time they could be perceived as a distant future research. It seems that this is an adequate place to say how thankful the author is to Prof. Rakowski's remarks and discussion on these studies.

2. Selected methods of artificial intelligence

2.1. Artificial neural networks

Artificial neural networks are computational systems that process information in ways similar to those that take place in the human brain [19, 23]. Algorithms based on neural networks may be used to model objects of unknown or intricate structure. Information they provide is of numeric character. Typical applications of neural networks include recognition, classification, comparative analysis, and object compression. A special application is the modeling of human perception and recognition processes which, among other issues, refer to sounds (e.g. speech recognition), music, audio signals, etc. Neural networks are widely discussed in numerous scientific monographs, and thus will not be presented in this work.

There are many methods of applying artificial neural networks to data processing. What differentiates them is: the model of neuron, the way of data propagation, and the algorithm of weights adaptation.

2.2. Rough set method

The concept of rough sets was introduced by PAWLAK, world renown scientist, at the beginning of 80s [12]. The rough sets method is mainly used to analyze data and to discover knowledge hidden in data. Main problems that find solutions in using rough sets are: data reduction (i.e. elimination of excessive information), discovering dependencies between data, identifying significant data, generating decisions upon data, performing approximate data classification, finding differences and similarities in data, recognizing and processing of image, identifying cause-result relations. The rough sets theory has many connections with numerous fields of applied informatics developed to explore issues of imprecision [1].

The theory introduces several characteristic definitions such as the *upper* and *lower approximation* or the *boundary region*. The concept of a rough set is based on the upper and lower approximations of sets. The main aim of applying rough sets theory is to divide the area of real data attributes into subsets corresponding to abstraction classes. The approximation is done by indicating the upper and lower approximations named respectively $\underline{B}X$ and $\overline{B}X$, where [12]:

$$\underline{B}X = \{x \mid [x]_B \subseteq X\}, \tag{1}$$

$$\overline{B}X = \{x \mid [x]_B \cap X \neq 0\}. \tag{2}$$

A rough set in D is defined as a family of all subsets U having the same lower and upper approximation.

Algorithms that employ the concept of rough sets theory are based on knowledge collection which is accomplished by saving specified attributes (parameters) that are then appropriately processed. The whole operation is performed to define the set of decision rules.

The set of data to be analyzed can be presented in the form of a table in which each single row represents a particular object. Attributes A describing the objects are stored in the columns of the table. The table organized in the described way is called an *information system* or a *Decision Table*. Formally an information system can be presented as [12]:

$$D = (U, A), \tag{3}$$

where U – non-empty, finite set of objects (*Universe*), A – non-empty, finite set of such attributes that, $a : U \to V_a$ for each $a \in A$. Set V_a is called a set of values a.

The decision table may be expressed as $D = (U, A \cup \{d\})$, where $d \notin A$. It may contain excessive data of two types: identical indiscernible objects and redundant attributes.

Algorithms employing rough sets may constitute grounds for an expert system. Such approach results in a set of IF ... THEN rules. At the same time, attributes used in a given rule constitute a so-called *reduct* which is a minimal set of parameters based on which a decision can be made [12]. It is possible to define a characteristic function called a rough measure for rules used in rough sets:

$$\mu_X^B(x) = \frac{\operatorname{card}(X \cap B(x))}{\operatorname{card} B(x)}. \tag{4}$$

Function $\mu_X^B(x)$ takes values from the range of [0, 1]. The value of the characteristic function is interpreted as the degree of element x membership to set X. The degree of membership can be considered as a conditional probability. The rules induced that have the measure $\mu_X^B(x) = 1$ are called certain rules (also known as deterministic, consistent, or compatible). The rules with rough measure of values less than 1 are called uncertain.

It can be easily seen that musical sounds or music analysis generates data that are approximate and uncertain, that is why rough sets are a valuable tool to deal with this inconsistency in data.

2.3. Fuzzy logic

The assumptions of fuzzy logic were formulated in 1965 by ZADEH [21]. The theory was then developed by many scientists, e.g. KANDEL, LEE, SUGENO, YAGER and others [18, 22]. It evolved in response to need of a means suitable to describe complex phenomena and loosely defined notions which a conventional mathematical apparatus could not describe appropriately. Classical methods of mathematics, based on a traditional concept of sets and a two-valued logic, are tools that confine the behavioral description of systems in which a human factor plays cardinal role. Such systems are present in many domains of today's life, e.g. economy, the theory of decision making, medicine and, as easily proved, in particular, perception of musical sounds music.

Generally speaking, fuzzy logic is a multi-valued logic. It is a device of artificial intelligence that bases on human-like reasoning. A fundamental property of fuzzy logic used in controlling systems is that it enables to define sets of linear and non-linear functions with intuitively formulated rules. This enables fuzzy logic to approximate human-like reasoning and interpretation of phenomena which, for majority of processes taking place in nature, is of continuous nature. The concepts of fuzzy logic are extensively described in literature [18, 22], thus only some background notions related to this theory have been presented in this work.

3. Automatic music information retrieval

The most important problems explored and shortly reviewed in the discussed experiments addressed issues of processing and classification of sounds produced by musical instruments with special regard to applications dealing with polyphonic sounds, musical phrases, human voice signals (singing), or meta descriptions of music (examples of areas of interest are shown in Fig. 1) [5–10, 20]. These issues constitute a new area of interests in musical informatics, that is the music information retrieval [24]. Since

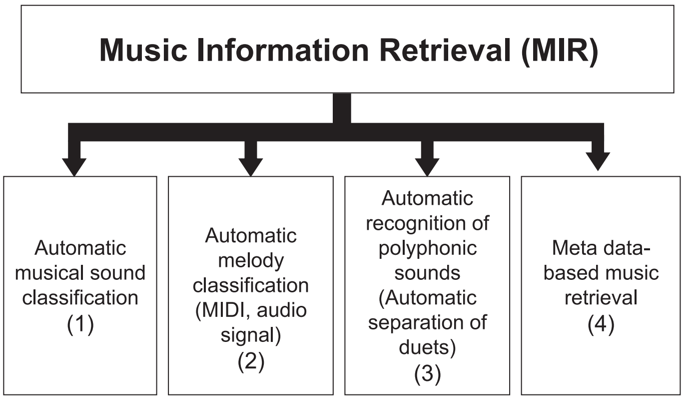


Fig. 1. Musical Information Retrieval.

the research carried out in this domain has a solid background now, many of the problems mentioned have been solved and found a way to applications in many domains, for example automatic classification of musical sounds or automatic search for music, to name a few. Also, MPEG 7 standard resulted from the research carried out in Music Information Retrieval domain [11].

It should be also mentioned that a school of musical informatics has been established at the Multimedia Systems Department resulting in many MSc. theses and Ph.D. studies in the MIR domain.

The process of musical data classification may in general be depicted as shown in a block diagram from Fig. 2. Queries issued to a decision system (formed as so-called *queries-by-example*) accept series of samples, symbolic musical notation, musical notation in XML, MIDI code, or meta description as information for the searches. As illustrated in Fig. 2, the main elements of the retrieval process are: music data acquisition (that in the case of musical sound signal comprises blocks of fundamental frequency detection and parametrization) and the decision system itself. The issue of parametrization has recently experienced extensive development which enforced the MPEG 7 standard definition. This standard includes a rich set of descriptors for music data parametrization [5, 11].

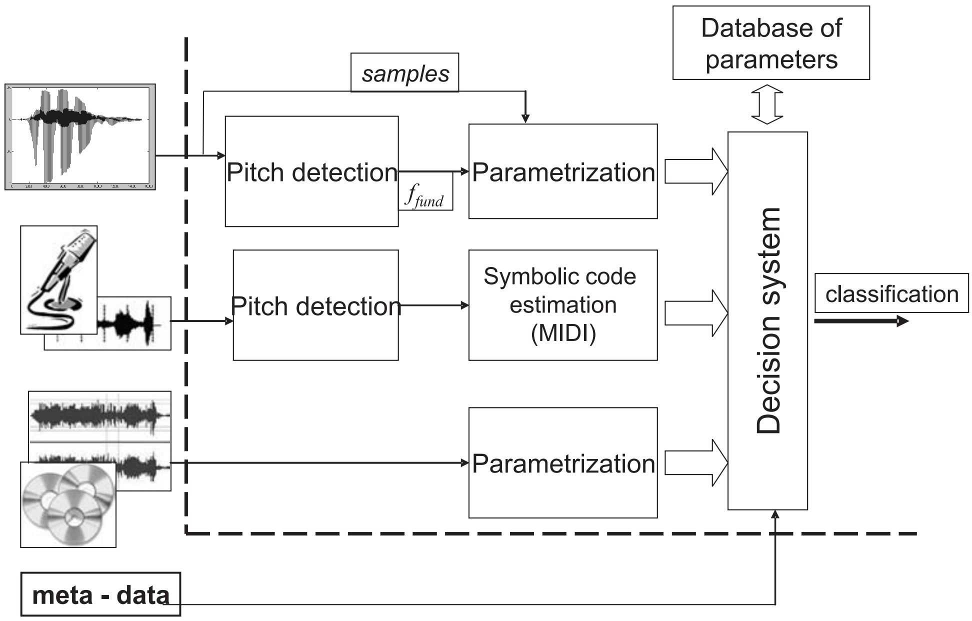


Fig. 2. Sample queries-by-example to the decision system.

The experiments conducted in the Multimedia Systems Department of the Gdańsk University of Technology that are shown here as an example of such query-by-example took as grounds 3500 samples of several musical instruments, and proved that the system is capable of classifying properly from 92.3% to 100% of sound samples depending on the number of categorized instruments and the algorithm used. It should be mentioned that musical sounds samples have been gathered from various musical databases, thus recorded in different acoustical environments. This was directed towards better generalization of a classifier. The results obtained for the classification of 10 musical instruments are shown in Table 1. The best achievements were produced by a group of neural networks (Table 2), and were about 1–2 percentage points better than those obtained for a single neural network (in both cases a number of classified instruments was 10). This was the effect of specializing each network to identify a smaller number of objects. Such approach allowed a 98% efficiency in recognition of instrument groups and classes for instruments with groups already appropriately classified. The resultant overall efficiency amounted to 96%.

Table 1. Comparison of classification efficiency for different algorithms.

Instrument	Single neural network			Group of neural networks			Rough sets method		
	total	errors	efficiency [%]	total	errors	efficiency [%]	total	errors	efficiency [%]
bassoon	137	5	96.35	146	5	96.58	112	4	96.4
clarinet	156	5	96.79	160	1	99.38	112	7	93.8
oboe	136	6	95.59	137	1	99.27	99	11	88.9
trombone	136	12	91.18	138	3	97.83	106	15	85.8
horn	144	13	90.97	123	21	82.93	98	12	87.8
saxophone	110	3	97.27	113	4	94.46	75	7	90.7
violin	140	3	97.86	141	18	87.23	116	11	90.5
trumpet	125	7	94.40	117	2	98.29	87	7	92.0
tuba	122	4	96.72	115	2	98.26	89	1	98.9
cello	125	9	92.80	141	4	97.16	126	7	94.4

Table 2. Efficiency of musical instruments group recognition.

Group of instruments	Number of sounds	Errors	efficiency [%]
string	411	6	98.54
woodwind	902	17	98.12
brass	712	18	97.47

Another issue explored was a widely understood analysis of musical phrases. The exploration scope embodied the search for an optimal vector of distinctive features for the melodic and rhythmic representations used to properly recognize musical styles and motifs, or to automatically generate the accompaniment. It turned out that in the case of musical style recognition the decision systems efficiency was worse than in the case of musical instrument sound classification. However, including additional selected parameters in meta descriptors induced improvement that allowed accurate (100%) efficacy.

Automatic search for specified motifs (started by the author of this paper and her collaborators) included the discovery of relations between units of music that constitute the hierarchy of structural arrangement in musical material [20]. The research, among other issues, aimed at the discovery of rhythmic patterns structure that would enable division into motifs, phrases, statements and musical periods in a way that corresponds to decomposition of a natural language sentence into hierarchical grammar structures. The discussed problem was related to the development of the methodology for finding an optimal vector of distinctive features necessary for melodic and rhythmic representations of musical materials with the use of neural network classifiers [5, 20]. Within the scope of these experiments the method to create all possible hierarchical rhythmic structures (also called rhythm hypotheses) was elaborated. The obtained hypotheses were then ranked with a decreasing function in order to determine which of them was the most appropriate rhythmic structure for a specific musical piece. The implemented algorithms are now applied to generate automated percussion background for melodic lines.

Undoubtedly, the most difficult problem to solve is the intelligent music retrieval from vast resources of the Internet databases. Besides methods based on artificial neural networks or Pawlak's rough sets, very interesting results for music search utilizing Pawlak's flow graphs were obtained [10]. The method was employed to search for musical pieces in the Internet resources based on information included in CD descriptions. Using descriptions of CD content or incomplete definitions given by users seeking information, the system should have determined the addresses of adequate files.

Neutralizing the effects of contextual discrepancy between file descriptions and the content of queries becomes possible, if algorithms based on methods invented to process natural languages, i.e. from the domain of *language engineering* are employed [10].

4. Fuzzy-logic pipe organ computer-controlled action

Classical pipe organs are certainly a special case in the domain of computer-controlled musical instruments. Such instruments are still frequently manufactured. Thus, quantitative and qualitative relations between the way pipe organs are controlled and the quality of this control, or what is even more important the sound fidelity of these instruments, must be clearly determined [3]. The bases of this work have been given by researchers in musical acoustics, e.g CADDY and POLLARD (1957) [2] and RAKOWSKI's papers from 1969 [14].

From the "musical" point of view, the way of controlling how pipes are activated must not only ensure that all particular sounds may be turned on, but it also has to guarantee musical articulation typical for pipe organs. Till now, these requirements were best satisfied by mechanical systems. The development of pipe organ control systems first led to pneumatic and then to electrical actions, which significantly facilitated playing these instruments. Unfortunately, as musicians say, actions of this type incorporate a new force between the key and the source of sound and deform the articulation to a meaningful degree. Additionally, instruments that use electrical actions react independently of the force with which the keyboard keys are pressed by the player and do not reflect musical articulation intended by the artist. The electromagnetic pipe valves open identically and thus do not allow to elicit nuances in tones, and just do not compare with classical organs having mechanical actions.

Differences between particular organs within the scope of various action systems directly influence how the instruments are estimated by musicians and listeners. So there was a special need to determine certain criteria, such as time- and spectrum-based parameters, which reflect phenomena occurring in a musical signal and are related to musical articulation, that would set standards on how organs are evaluated and to relate these criteria to subjective estimation. The research also considered using fuzzy logic-based solutions in pipe organ action systems. For this purpose, a pipe organ control system using fuzzy logic was designed and constructed (Fig. 3) [3, 4, 7].

Controlling the movement of an electromagnetic valve is a very difficult task, as the velocity of the electromagnet depends non-linearly on the activating current. Moreover, the sensor of the key motion located under the keyboard does not linearly reflect the way the key is pressed. Precise description of these non-linear functions is very difficult, or even impossible. This is where fuzzy logic brings solution to the problem.

The constructed system analyzes two input values (see Fig. 3a), that is: the velocity with which the key is depressed and the key number. The fuzzy-logic-based controller must determine the number of the key since the process of sound rise depends on the dimensions of pipes, and thus is different for 8’, 16’ and larger pipes which are mapped to the “lower” part of the keyboard, than for small pipes corresponding to the keys from the “upper” part of the keyboard. The values that control the system of pipes stimulation, i.e. the velocity and the number of the key, are obtained from a synthesizer which is equipped with a dynamic keyboard and generates these values encoded in MIDI (see Fig. 3b). The software created based on the rules of fuzzy logic processes specified input parameters in real time and converts them into output values that determine how fast the pipe valve will be opened. The implementation of the system employs the circuit of electromagnets in which the value of activating current influences the resultant rate with which the pipes are supplied with air.

Examples of analyses of the time- and frequency-domain characteristics of the recorded sounds from the constructed system are presented in Figs. 4 and 5.

The plots show the differences that are visible in the time representation of the analyzed sounds, as well as in the representation of waterfall plots, respectively for fast (Figs. 4a and 5a) and slow (Figs. 4b and 5b) opening of the valve.

a)

b)

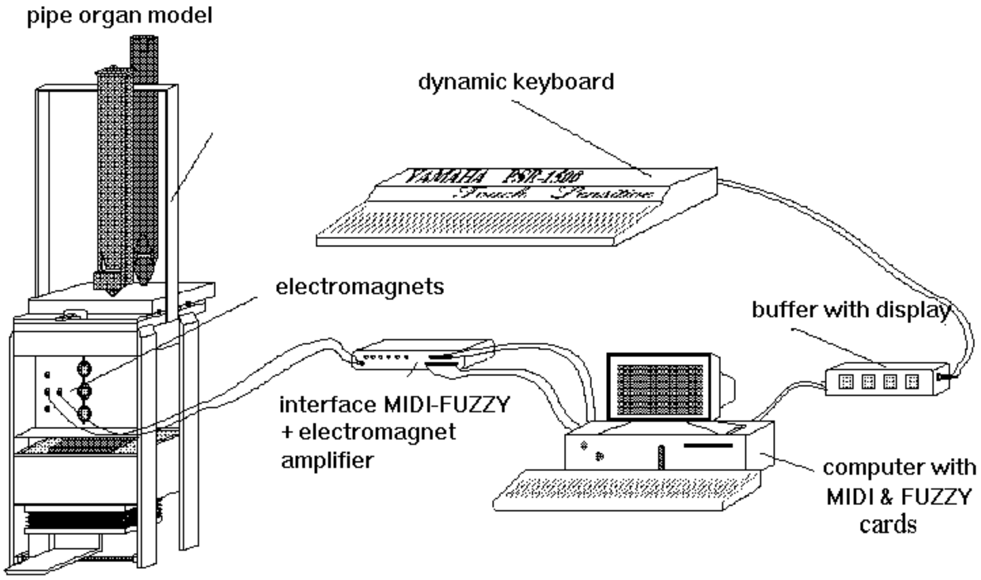
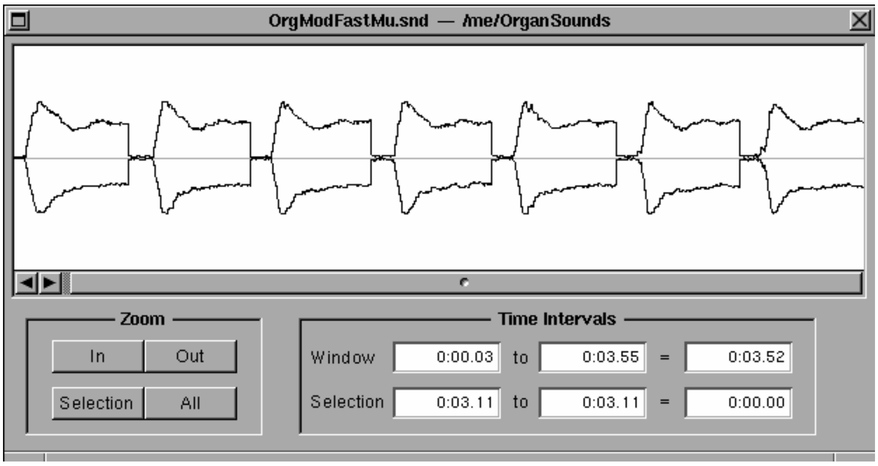


Fig. 3. Model of pipe organ (a), control system for pipe organ action (b) [7].

a)



b)

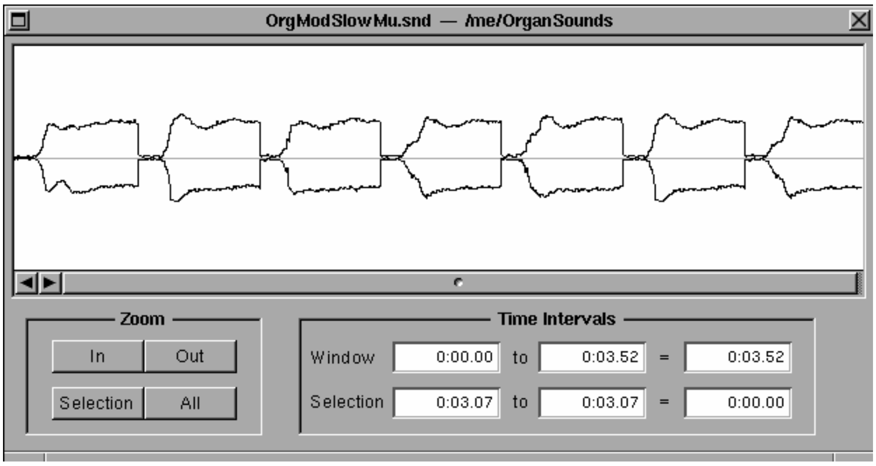
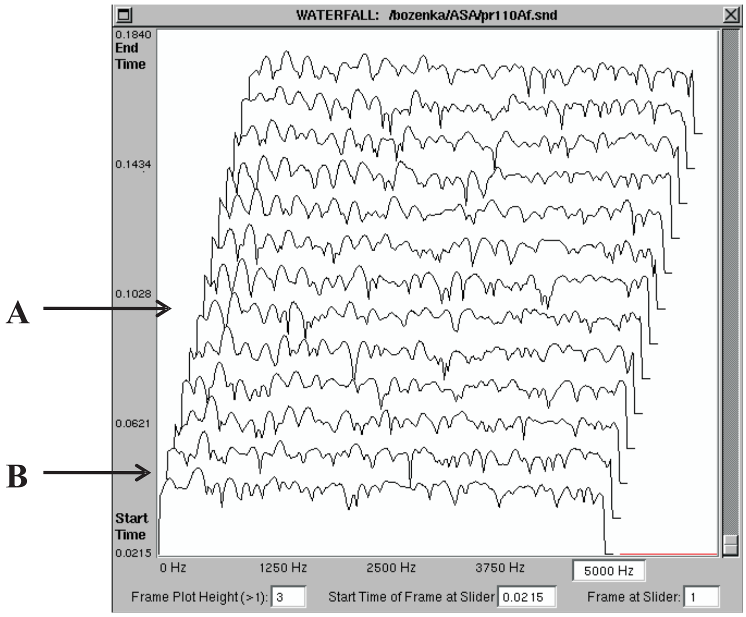


Fig. 4. Analyses of time-domain characteristics of the sounds of Principal 8’ in the case of: fast opening of the valve (a), slow opening of the valve (b).

Both spectral characteristics differ mainly in the behavior of the second harmonic whose dynamic of change depends directly on the rate at which the key is depressed – the faster the key is depressed, the quicker the second harmonic grows. There are also other discrepancies. It is easy to observe that the fundamental is much weaker when the key is depressed quickly. Arrows “A” in Fig. 5 show the starting point of the rising of fundamental, whereas arrows ‘B’ show the rising of the second harmonics.

Therefore, it may be said that the constructed fuzzy logic control system for an electric pipe organ action responds properly depending on differentiated musical articulation, enriching music by providing nuances to its interpretation.

a)



b)

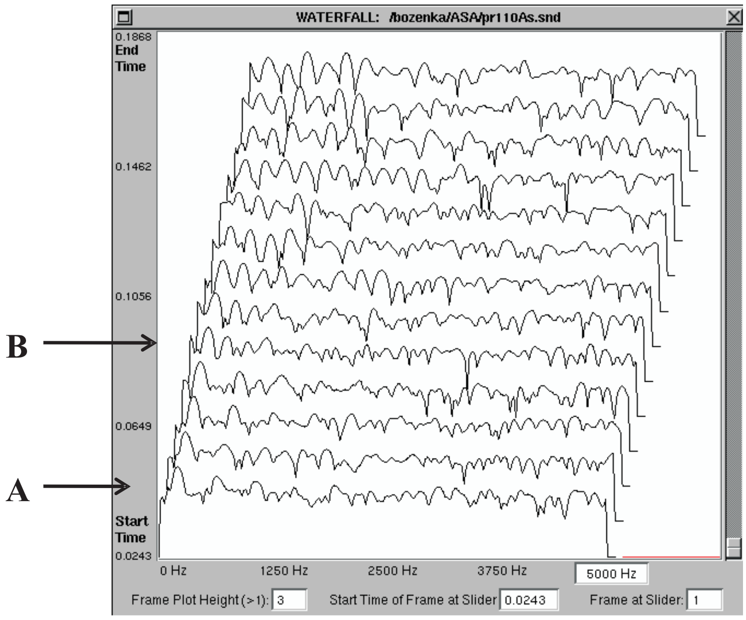


Fig. 5. Analyses of frequency-domain characteristics of the sounds of Principal 8’ in the case of: fast opening of the valve (a), slow opening of the valve (b).

5. Summary

The research conducted in the Multimedia Systems Department at the Gdańsk University of Technology, which comprised applications of the selected soft computing methods in the domain of musical acoustics, particularly in classification and recognition of musical sounds led to a novel approach to issues of audio data retrieval and classification. These issues, due to their complexity and incomplete recurrence, evade the analyses based on deterministic models. Thus, the use of methods originating from soft computing seems fully justified when solving essential problems from the domain of widely understood sound engineering or acoustics, especially musical acoustics. Among others, the research proved that fuzzy logic-based electrical organ action may be successfully implemented in the control system of pipe organs.

References

[1] BAZAN J. G., NGUYEN H. S., SKOWRON A., SZCZUKA M., *A view on rough set concept approximations*, [in:] WANG G., LIU Q., YAO Y., SKOWRON A. [Eds.], Proc of RSFD. Lecture Notes in Computer Science 2639, Springer, Chongqing, 181–188, 2003.

[2] CADDY S., POLLARD H. F., *Transient sounds in organ pipes*, Acustica, **7,** 227–280 (1957).

[3] KOSTEK B., *Untersuchungen an Orgeltrakturen unter dem Aspekt musikalischer Artikulierung*, [in:] Fortschritte Der Akustik, Teil A, DAGA '92, Berlin, 245-248, 1992.

[4] KOSTEK B., *Articulation-related features in the pipe organ sound*, Archives of Acoustics, **22,** 219–244 (1997).

[5] KOSTEK B., CZYŻEWSKI A., *Representing musical instrument sounds for their automatic classification*, J. Audio Eng. Soc., **49,** 768–785 (2001).

[6] KOSTEK B., *Soft Computing in acoustics, applications of neural networks, fuzzy logic and rough sets to musical acoustics*, Physica Verlag, Heidelberg, New York 1999.

[7] KOSTEK B., *Perception-based data processing in acoustics. Applications to music information retrieval and psychophysiology*, Springer Verlag, Studies in Computational Intelligence, Berlin, Heidelberg, New York 2005.

[8] KOSTEK B., *Application of soft computing to automatic music information retrieval*, J. American Society for Information Science and Technology, **55,** 12, 1108–1116 (2004).

[9] KOSTEK B., *Musical instrument classification and duet analysis employing music information retrieval techniques*, Proc. of the IEEE, **92,** 4, 712–729 (2004).

[10] KOSTEK B., CZYŻEWSKI A., *Processing of musical metadata employing Pawlak's flow graphs, Rough Set Theory and Applications (RSTA), Transactions on rough sets, advances in rough sets*, LNCS 3100, **1,** 285–305 (2004).

[11] LINDSAY A. T., HERRE J., *MPEG-7 and MPEG-7 audio – an overview*, J. Audio Eng. Soc., **49,** 7/8, 589–594 (2001).

[12] PAWLAK Z., *Rough sets*, J. Computer and Information Science, **11**, 5 (1982).

[13] RAKOWSKI A., *Opening transients in tones of the flute*, Bull. Soc. Ami. Sc. et Lt., Poznań, B19, 147–161 (1976).

[14] RAKOWSKI A., *Eine Analyse des Intonierungsvorganges bei Orgeln*, Gravesaner Blätter, **15,** 46–58 (1969).

[15] RAKOWSKI A., *The application of organ sound in reverberation measurements in the concert hall of Warsaw Academy of Music*, Archives of Acoustics, **2,** 105–113 (1977).

[16] RAKOWSKI A., *Measurements of pitch*, J. of the Catgut Acoustical Society, **27,** 11–29 (1977).

[17] RAKOWSKI A., *Studies on musical pitch and timbre* [in Polish], F. Chopin Music Academy, Warsaw 1999.

[18] TAKAGI T., SUGENO M., *Fuzzy identification of systems and its application to modelling and control*, IEEE Trans on Systems, Man and Cybernetics, **15,** 116–132 (1985).

[19] TADEUSIEWICZ R., *Sztuczne sieci neuronowe* [in Polish], Akademicka Oficyna Wydawnicza RM, Warsaw 1993.

[20] WÓJCIK J., KOSTEK B., *Intelligent methods for musical rhythm finding systems*, [in:] Intelligent Technologies for Inconsistent Knowledge Processing (NGUYEN N. T., [Ed.]), **10,** 11, 187–202, (2004).

[21] ZADEH L., *Fuzzy sets*, Information and Control, **8,** 1965, 338–353 (1965).

[22] ZADEH L., KACPRZYK J. [Eds.], *Fuzzy logic for the management of uncertainty*, Wiley, New York 1992.

[23] ŻURADA J., *Introduction to artificial neural systems*, West Publishing Comp., St. Paul, 1992.

[24] URL: http://ismir2006.ismir.net/ Music Information Retrieval website, 2006.

[25] URL: http://www.soft-computing.de/def.html (A Definition of Soft Computing – adapted from L.A. Zadeh), 2007.