Applications of Rough Sets and Neural Nets to Noisy Audio Enhancement

Rafal Krolikowski, Andrzej Czyzewski
Sound Engineering Department, Technical University of Gdansk
ul. Narutowicza 11/12, 80-952 Gdansk, Poland
Phone: +48 -58 3472301, Fax: +48 -58 3471114
e-mail: {rafal, andrzej}@sound.eti.pg.gda.pl

ABSTRACT: A new concept of reduction of non-stationary noise affecting audio signals transmitted in telecommunication channels is proposed. This concept exploits some features of the human auditory system as well as some methods originated from artificial intelligence domain, i.e. reasoning based on rough sets and neural processing.
The foundations of the engineered method together with a description of applied intelligent decision algorithms are presented in the paper. A number of experiments has been prepared, and some of them have already been carried out. Hence, a brief discussion of the results of these experiments and some conclusions are also included in the paper. The main focus is put on a comparison of different intelligent methods used to non-stationary noise reduction.

KEYWORDS: reduction of non-stationary noise; rough sets; neural networks; perceptual coding of audio

INTRODUCTION

The commonly used noise reduction methods do not use some subjective properties of the human auditory system, which have been successfully exploited in some audio coding standards. However, as was revealed by the results of experiments carried out by the authors, auditory masking can be also used to the suppression of noise corrupting audio signals. The mathematical foundations of this perceptual approach and relevant algorithms were presented in some recent authors' papers (Czyzewski 1999) (Krolikowski 1999).
In all noise reduction methods, there is a need to know at least approximated statistics of the noise. This problem becomes more complex in the case of non-stationary noise, since such a method requires to choose a certain noise statistics from among others. Hence, a problem of an efficient decision system occurs. Therefore intelligent algorithms (i.e. rough sets or neural networks) can be very helpful as decision systems (Czyzewski 1997a) (Czyzewski 1997b).

BASIC PRINCIPLES OF PSYCHOACOUSTICS

The concept of critical bands is related to propagation and processing of acoustic signals in the human auditory system. Well-proven phenomena reveal that the inner ear behaves as a bank of band-pass filters which analyse a broad spectral range in subbands independently from others. These subbands are called critical bands, and a perceptual unit of frequency has been introduced. It is called Bark and is related to the width of a single subband. Often used transformation to this subjective scale of hearing is the following relation proposed by Zwicker (Zwicker 1990):

$$b = 13 \cdot \mathrm{arctg}\left(0.76 \cdot 10^{-3} \cdot f\right) + 3.5 \cdot \mathrm{arctg}\left[\left(f / 7500\right)^2\right], \tag{1}$$

where b, f denote frequency in Barks and Hz, respectively.
Another psychoacoustic phenomenon is related to masking. Some tones can be inaudible in the presence of others, especially when one of them is louder, and their frequencies are not too distant. These tones which mask others are called *maskers*, and this phenomenon is fundamental for contemporary audio coding standards (Shlien 1994). More detailed information on psychoacoustics can be found in abundant literature (Zwicker 1990).

DESCRIPTION OF THE PERCEPTUAL NOISE REDUCTION SYSTEM

The perceptual noise reduction system (Fig. 2) is fed by two inputs: the noisy signal $y(m)$ and the noise patterns $\tilde{n}(m)$. The signal $y(m)$ consists of the original audio signal $x(m)$ corrupted by the noise $n(m)$, and is transformed to the spectral representation $Y(\mathrm{j}w)$ with the use of the DFT procedure. In turn, the patterns $\tilde{n}(m)$ are assumed to be correlated to the noise $n(m)$, and are taken from empty passages of the signal transmitted in a telecommunication channel. The signal $\tilde{n}(m)$ is delivered to the Noise Estimation Module which task is to collect essential information on the noise $n(m)$. At its output, the time-frequency noise estimation $r(t,\mathrm{j}w)$ is obtained. Both this estimation $r(t,\mathrm{j}w)$ and the spectrum of the corrupted audio $Y(\mathrm{j}w)$ are supplied to the Decision Systems. Its first task is to select one of the collected spectral estimations $r(\mathrm{j}w) \subset r(t,\mathrm{j}w)$ which is correlated best to the corrupting noise in a given moment. The second task is to qualify the elements of the signal $Y(\mathrm{j}w)$ for two disjoint sets: the set *U* of the useful or the set *D* of the useless elements. It is necessary to know, which spectral components are maskers (useful), and which ones are to be masked (useless). Thus, in the case of the use of the *N*-point DFT the following condition is fulfilled:

$$N/2 = |U| + |D|, \tag{2}$$

where $|U|$, $|D|$ are the numbers of useful and useless elements.

The spectrum of the corrupted signal $Y(\mathrm{j}w)$ as well as the sets *U* , *D* and the chosen noise estimation $r(t,\mathrm{j}w)$ are fed to the Perceptual Noise Reduction Module that executes a perceptual algorithm of noise reduction. Next, the output $\hat{Y}(\mathrm{j}w)$ is processed by the inverse DFT procedure, and finally the restored signal $\hat{y}(m)$ is obtained, which is subjectively perceived as less noisy than the original one.

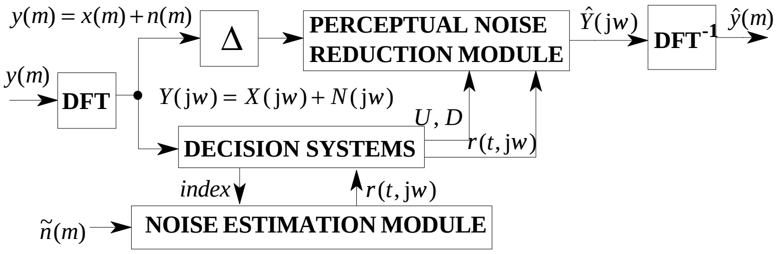

Figure 1: General lay-out of the noise reduction system.

IMPLEMENTATION OF THE PERCEPTUAL NOISE REDUCTION SYSTEM

NOISE ESTIMATION MODULE

As was stated, the task of the Noise Estimation Module is to collect information on the noise $n(m)$ corrupting audio on the basis of analysis of the noise patterns $\tilde{n}(m)$. The module's run can be divided into two modes: the noise analysis mode and the noise reduction mode, as was illustrated in Fig. 2. In the first one (Fig. 2a), the patterns $\tilde{n}(m)$ are analysed in the *Extraction of Noise Parameters* block where they are transformed into the spectral domain, averaged upon subsequent *L* frames and analysed. As a result, two kinds of output are obtained: the average power spectrum $\hat{\boldsymbol{N}}_k$ which can be referred to as a vector of spectral power values for consecutive frequency components, and the associated vector $\boldsymbol{V}_k^{\hat{\boldsymbol{n}}}$ of coefficients related to the spectrum $\hat{\boldsymbol{N}}_k$. The index *k* denotes the time interval, within which elements of these vectors are computed. Subsequent, the both vectors are collected in the *Table of Vectors*. Thus, this table can 'remember' changes of the noise $\tilde{n}(m)$ statistics within time. The content of the *Table of Vectors* is used during the training of decision algorithms in the Decision Systems module and the noise reduction mode (Fig. 2b). In this latter mode, due to a query *index* to the table, the appropriate spectrum $\hat{\boldsymbol{N}}_j$ is output, which is expected to be most correlated to the noise currently corrupting the audio. The query *index* value is produced by a decision system, and denotes the index of a desired spectrum in the table.

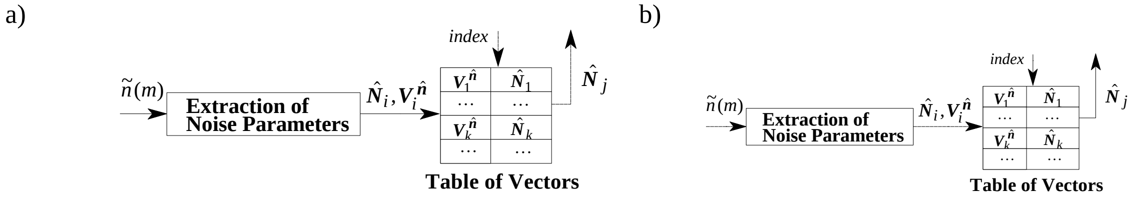

Figure 2: Detailed scheme of the Noise Estimation Module: (a) noise analysis mode, (b) noise reduction mode. Dashed lines denote inactive connections.

In the case of use of the *N*-point DFT the vector $\hat{\boldsymbol{N}}_k$ is defined as below:

$$\hat{\boldsymbol{N}}_k = \left[\hat{N}_{1,k} \ \ldots \ \hat{N}_{n,k} \ \ldots \ \hat{N}_{N/2,k}\right]^T , \tag{3}$$

where the *n*-th element $\hat{N}_{n,k}$ is averaged on the basis of last *L* values of the spectral power $\tilde{N}_n$ according to the following formula:

$$\hat{N}_{n,k} = \frac{1}{L} \cdot \sum_{l=(k-1)\cdot L+1}^{k\cdot L} \tilde{N}_n^{(l)} \tag{4}$$

The associated vector $\boldsymbol{V}_k^{\hat{n}}$ serves as a key vector and is exploited during the noise reduction mode when a spectrum $\hat{\boldsymbol{N}}_j$ is searched for (Fig. 2b). This vector should be unique, however in practice the condition is hard to be ensured. Its elements are expected to reflect quantitatively a noisy character of the average spectrum $\hat{\boldsymbol{N}}_k$. Therefore two kinds of parameters are considered that turned out to be very robust in contemporary perceptual coding schemes (Shlien 1994): the *Spectral Flatness Measure* (Johnston 1988) and the *unpredictability measure* (Brandenburg 1990). These parameters are computed in each critical band, and their definitions for the *l*-th frame are given below.

- Application of Spectral Flatness Measure

The *SFM* parameter is defined as the ratio of the geometric to the arithmetic mean of the power spectrum (Johnston 1998), and is expressed in dB. In the *b*-th critical band, the parameter can be redefined as follows:

$$SFM_b^{(l)} = 10 \cdot \log_{10} \frac{G_m^{(l)}}{A_m^{(l)}} = 10 \cdot \log_{10} \frac{\left[\prod_{i=lower(b)}^{upper(b)} S_i^{(l)}\right]^{1/count(b)}}{\frac{1}{count(b)} \cdot \sum_{i=lower(b)}^{upper(b)} S_i^{(l)}} , \tag{5}$$

where S_i is the spectral power of the *i*-th frequency component, which is obtained by means of the *N*-point short-time Discrete Fourier Transform, whereas *lower*(*b*) and *upper*(*b*) denote indexes of the first and the last spectral component in the *b*-th subband which contains *count*(*b*) components.
Hence, the vector $\boldsymbol{V}_k^{\hat{n}}$ can be described in the following way:

$$\boldsymbol{V}_k^{\hat{n}} = \left[SFM_{1,k} \ \ldots \ SFM_{b,k} \ \ldots \ SFM_{B,k}\right]^T , \quad \text{where:} \quad SFM_{b,k} = \frac{1}{L} \cdot \sum_{l=(k-1)\cdot L+1}^{k\cdot L} SFM_b^{(l)} \tag{6}$$

- Application of the unpredictability measure

Introducing denotations of the spectral magnitude prediction $\hat{r}_i^{(l)}$ and the phase prediction $\hat{f}_i^{(l)}$ of the *i*-th spectral component on the basis of the their last two real values as below:

$$\begin{cases} \hat{r}_i^{(l)} = r_i^{(l-1)} + (r_i^{(l-1)} - r_i^{(l-2)}) \\ \hat{f}_i^{(l)} = f_i^{(l-1)} + (f_i^{(l-1)} - f_i^{(l-2)}) \end{cases} \Rightarrow \begin{cases} \hat{r}_i^{(l)} = 2 \cdot r_i^{(l-1)} - r_i^{(l-2)} \\ \hat{f}_i^{(l)} = 2 \cdot f_i^{(l-1)} - f_i^{(l-2)} \end{cases} , \tag{7}$$

the unpredictability measure $c_i^{(l)}$ is defined as the Euclidean distance between the real values of $r_i^{(l)}$, $f_i^{(l)}$ and the predicted ones of $\hat{r}_i^{(l)}$, $\hat{f}_i^{(l)}$ according to the formula (Brandenburg 1990):

$$c_i^{(l)} = \frac{\mathrm{dist}\left((\hat{r}_i^{(l)}, \hat{f}_i^{(l)}), (r_i^{(l)}, f_i^{(l)})\right)}{r_i^{(l)} + \mathrm{abs}\left(\hat{r}_i^{(l)}\right)} = \frac{\sqrt{\left(r_i^{(l)} \cdot \cos f_i^{(l)} - \hat{r}_i^{(l)} \cdot \cos \hat{f}_i^{(l)}\right)^2 + \left(r_i^{(l)} \cdot \sin f_i^{(l)} - \hat{r}_i^{(l)} \cdot \sin \hat{f}_i^{(l)}\right)^2}}{r_i^{(l)} + \left|\hat{r}_i^{(l)}\right|} \tag{8}$$

In such a case, the vector $\boldsymbol{V}_k^{\hat{\boldsymbol{n}}}$ can described as below:

$$\boldsymbol{V}_k^{\hat{\boldsymbol{n}}} = \left[C_{1,k} \ \dots \ C_{b,k} \ \dots \ C_{B,k}\right]^T, \tag{9}$$

where the element $C_{b,k}$ is calculated for the b-th critical band and averaged upon last L frames in the following way:

$$C_{b,k} = \frac{1}{L} \cdot \sum_{l=(k-1)\cdot L+1}^{k \cdot L} C_b^{(l)}, \quad \text{where: } C_b^{(l)} = \frac{1}{count(b)} \cdot \sum_{i=lower(b)}^{upper(b)} c_i^{(l)} \tag{10}$$

DECISION SYSTEMS

As was mentioned, the respective objective of the Decision System module are as follows: selection of the noise spectrum $\hat{\boldsymbol{N}}_j$ that best matches the corrupting noise in a given moment and qualification of the spectral components for the sets of useful and useless elements.

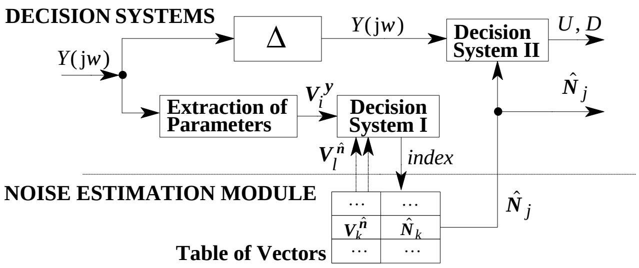

Figure 3: Detailed scheme of the Decision Systems.

The module is fed by the spectral representation of the noisy signal $Y(jw)$. First, the input signal is processed in the *Extraction of Parameters* block which task is to obtain a vector of parameters $\boldsymbol{V}_i^{\boldsymbol{y}}$, and these parameters are expected to be mostly related to the noisy character of the input $Y(jw)$. Therefore elements of this vector are defined by analogy to the key vector $\boldsymbol{V}_k^{\hat{\boldsymbol{n}}}$ elements, and hence computed as in the formulae (5)-(10). The vector $\boldsymbol{V}_i^{\boldsymbol{y}}$ is next supplied to the *Decision System* I which objective is to give the *index* value of the noise spectrum $\hat{\boldsymbol{N}}_j$ that should be most correlated to the noise present in the noisy signal $Y(jw)$. Having received the desired vector $\hat{\boldsymbol{N}}_j$ from the *Table of Vectors*, this vector is compared with the spectral representation $Y(jw)$ in the *Decision System* II which produces two output sets: the set U of useful and the set D of useless components.

Implementation of the Decision System I

Making decision in the *Decision System* I can be based on the rough set method or neural reasoning, and the work of the system can be divided into two modes: the training mode and the execution mode. In the first case, the content of the *Table of Vectors* is exploited, which is depicted in Fig. 3 with dashed lines. The intelligent decision algorithms are considered further in the paragraph.

- Application of rough sets

In the training mode, related to rule discovery, a part of the *Tables of Vectors* is treated as a decision table, where elements of the key vector $\boldsymbol{V}_k^{\hat{\boldsymbol{n}}}$ defined by the formulae (3) or (6) are conditional attributes and the vector's index in the *Table of Vectors* is a decision attribute. Therefore the *k*-th object in the table is described by the following relation:

$$SFM_{1,k}\ ,\ldots\ ,SFM_{b,k}\ ,\ldots\ SFM_{B,k} \quad \Rightarrow \quad k \qquad \text{or} \qquad C_{1,k}\ ,\ldots\ ,C_{b,k}\ ,\ldots\ C_{B,k} \quad \Rightarrow \quad k\ , \tag{11}$$

where the parameters: *SFM*, *C* are computed according to the expressions (6) and (10), respectively.
The rule discovery procedure is based on the rough set principles (Pawlak 1982) and algorithm (Czyzewski 1997b). It can be noticed that only conditional attributes require quantization. So far, only the uniform quantization is proposed.
In the execution mode, the input vector of noisy audio parameters $\boldsymbol{V}_i^{\boldsymbol{y}}$ is quantized, and next processed by the set of generated rules. In result, the *index* value of the noisy spectrum $\hat{\boldsymbol{N}}_j$ in the *Tables of Vectors* is obtained.

- Application of neural networks

The applied neural network is a classic feedforward structure with hidden neurons (Zurada 1992). In the preliminary experiments, only one hidden layer was considered. The number of input neurons is equal to the number of elements in the key vector $\boldsymbol{V}_k^{\hat{\boldsymbol{n}}}$, and these elements are given by the formulae (3) and (6). There is only one output neuron which produces the entry value (*index* value) to the *Table of Vectors*. For all input and hidden units, the neuron's activation function is sigmoidal. However, since the index *value* is an integer, this function for the output neuron can is considered as sigmoidal or linear.
In the training mode, only a part of the *Table of Vectors* is used in the training set: the key vector $\boldsymbol{V}_k^{\hat{\boldsymbol{n}}}$ is an input vector, whereas its index in the table is a desired output, associated with $\boldsymbol{V}_k^{\hat{\boldsymbol{n}}}$. Thus, the training set is as follows:

$$\left\{(\boldsymbol{V}_1^{\hat{\boldsymbol{n}}},1),\ldots,(\boldsymbol{V}_k^{\hat{\boldsymbol{n}}},k),\ldots,(\boldsymbol{V}_K^{\hat{\boldsymbol{n}}},K)\right\} \tag{12}$$

As a training method, the standard Error Backpropagation Training Algorithm (Zurada 1992) is applied, and the error measure is based on the mean squared error. If the neuron's activation function of the output neuron is sigmoidal, the desired index value should be scaled down in order to match the interval [0,1].
In the execution mode, the network's output (related to the index value) must be rounded up to the nearest integer. Additionally, if the output unit processes according to sigmoidal function, before the round-up the neuron's output must be scaled up by the same factor by which was scaled down.

Implementation of the Decision System II

In the *Decision System* II, the division into useful and useless elements is executed according to the following simple procedure. All these components which spectral powers *Y* exceed the double average value of the representative noise estimation $\hat{\boldsymbol{N}}_j$ are assumed to be the useful elements. In turn, the remaining components are regarded as useless ones. Hence, in the case of use of the *N*-point DFT the sets *U* and *D* can defined as follows:

$$\begin{aligned} U &= \left\{n, Y_n: \quad Y_n \geq 2\cdot\hat{N}_{n,j} \quad \text{and} \quad n = 1,\ldots,N/2\right\} \\ D &= \left\{n, Y_n: \quad Y_n < 2\cdot\hat{N}_{n,j} \quad \text{and} \quad n = 1,\ldots,N/2\right\} \end{aligned} \tag{13}$$

In general, there can be many various methods of such a division (Czyzewski 1997a, Czyzewski 1997b), and a choice of one of them has significant influence on the subjective quality of a restored audio signal.

PERCEPTUAL NOISE REDUCTION MODULE

The task of the module is to process the spectral representation of the noisy signal as follows. All useful spectral components are reduced according to the spectral subtraction principles (Vaseghi 1997), whereas the remaining useless components are masked using the psychoacoustic approach. This perceptual approach is a separate complex issue which is not related to the application of intelligent tools, and due to the space limitations it is not described in this paper. However, the applied perceptual models, engineered methods together with appropriate algorithms were presented extensively in details in the recent publications of the authors (Czyzewski 1999) (Krolikowski 1999).

EXPERIMENTS

There were two objectives of the experiments: verification of the engineered method for non-stationary noise reduction and comparison of different decision algorithms. Some verification tests were carried out first, in order to check, whether application of intelligent tools could improve the quality of restored audio signals. The results were encouraging enough, so that the next comparison experiments were prepared.

In order to assess the quality of a decision algorithm, the result of such an algorithm should be compared with the desired output. In the case of the research, it was necessary to know whether the noise spectrum pointed by the decision system was the best choice, and if not, an error measure was needed.

For the purpose of the comparison experiments, two recordings were made: a male voice and a non-stationary noise taken from a radio channel. Next, the original audio was corrupted by the additive noise, and in the same time, elements of the key vectors of the noise were computed and collected. Since the original audio and the noise were given, it was known, which part of the noisy voice was described by which key vector and noise spectrum vector.

The parameters of these recordings were as follows. They both were mono, sampled with 16-bit resolution and with the frequency equalled to 8192 Hz, which resulted in B = 18 critical bands. Their length was 5.81 s for the voice signal and 2.79 s for the noise.

During the audio processing, the signals were divided into frames and overlapped. Since the Hamming window function was used, the overlap size was the half of the frame length N. In the experiments three values of N are considered: N = 128, N = 256 and N = 512, and their influence on the time- and frequency resolution are shown in Tab. 1. Additionally, in Tab. 1 there is a number of the noise key vectors, assuming that signal frames are averaged upon last L = 4 frames. It can be notices that the number of the key vectors is also the number of objects in a decision table or the number of training vectors.

Table 1: Influence of the frame size N on training set and the time- and frequency resolution.

N	***time-resolution***	***frequency resolution***	***number the training vectors***
128	7.83 ms	64 Hz	88
256	15.63 ms	32 Hz	44
512	31.25 ms	16 Hz	22

The comparison tests were divided with respect to the following variable parameters:

- various frame size: N = 128, N = 256, N = 512
- various key vector types: based on the *SFM* parameters (6) and the unpredictability measure C (9)
- various quantization steps: 0.1, 0.5, 1
- various number of hidden neurons: 10, 15, 20
- various neuron's activation function of the last unit: sigmoidal, linear

Hence, a single survey can be described by a set of parameters which are valid for a given decision algorithm. Thus, the following denotations are proposed: (N, *vector*, RS, *quantization step*) for application of rough sets, and (N, *vector*, NN, *hidden neurons*, *output unit*) for application of neural nets. Totally, there are 32 surveys. So far, only three of them have been completed. They are: (512,*SFM*,NN,10,sigmoidal), (512,*C*,NN,10,sigmoidal), (512,*SFM*,RS,0.5).

In order to assess the quality of a decision system, the error measure E is introduced. This measure is the average value of the errors $E^{(i)}$ produced for all I frames of the noisy signal, and is defined as follows:

$$E = \frac{1}{I} \cdot \sum_{i=1}^{I} E^{(i)} \quad \text{and} \quad E^{(i)} = \sum_{b=1}^{B} \left(V_{b,i}^{y} - V_{b,index}^{\hat{n}} \right)^2 , \tag{14}$$

where $V_{b,i}^{y}$ is the b-th element of the vector of parameters of the noisy audio, whereas $V_{b,index}^{\hat{n}}$ is the b-th element of the key vector which is placed at position *index* in the *Table of Vectors*.

By analogy, the optimal error measure E_{opt} can be introduced. Assuming that the i-th frame of the audio is corrupted by the noise described by the j-th vector in the *Table of Vectors*, this measure is defined as below:

$$E_{\text{opt}} = \frac{1}{I} \cdot \sum_{i=1}^{I} E_{\text{opt}}^{(i)} \quad \text{and} \quad E_{\text{opt}}^{(i)} = \sum_{b=1}^{B} \left(V_{b,i}^{y} - V_{b,j}^{\hat{n}} \right)^2 \tag{15}$$

Additionally, the maximum error $E_{\max}$ which can occur is proposed. This measure is given by the following formula:

$$E_{\max} = \frac{1}{I} \cdot \sum_{i=1}^{I} E_{\max}^{(i)} \quad \text{and} \quad E_{\max}^{(i)} = \max_{k=1,\ldots,K} \left[\sum_{b=1}^{B} \left(V_{b,i}^{y} - V_{b,k}^{\hat{n}} \right)^2 \right], \tag{16}$$

where K is the number of vectors in the *Table of Vectors*.
Finally, the quality coefficient q is introduced, which is calculated according to the following expression:

$$q = 1 - \frac{\left| E - E_{\mathrm{opt}} \right|}{E_{\max} - E_{\mathrm{opt}}} \tag{17}$$

For the completed surveys this quality coefficient is as follows:

- (512,*SFM*,NN,10,sigmoidal): q = 81.38 %
- (512,*C*,NN,10,sigmoidal): q = 85.21 %
- (512,*SFM*,RS,0.5): q = 78.43 %

It can be assumed that the best result for the survey (512,*C*,NN,10,sigmoidal) is due to the use of the more precise key vector type (based on the unpredictability measure). In turn, the worst result for rough sets can be caused mainly by the inefficient quantization step.

CONCLUSIONS

In the paper, the engineered system for non-stationary noise reduction has been presented, which exploits reasoning based on neural processing and rough sets. A number of tests has been planned and carried out. Their results are encouraging and suggest that intelligent algorithms can support making decision in the engineered noise reduction system. The further experiments are expected to answer the following questions and doubts:
- whether (and to which extend, if necessary) parameters describing a noisy character of an audio signal are efficient ?
- whether advanced decision systems such as neural networks or rough sets are needed ?
- which decision system (and for which parameters) is the most efficient in the presented system ?

REFERENCES

Brandenburg K., 1990, ”Second Generation Perceptual Audio Coding: The Hybrid Coder”, 90th Audio Engineering Society Convention, Monteux, France, preprint 2937.

Czyzewski A., 1997, ”Learning Algorithms for Audio Signal Enhancement.”, Journal of Audio Engineering Society, vol. 45, pp. 931-943.

Czyzewski A., Krolikowski R., 1997, „Application of Intelligent Decision Systems to the Perceptual Noise Reduction of Audio Signals”, 5th European Congress on Intelligent Techniques and Soft Computing, Aachen, Germany, vol. 1, pp. 188-192.

Czyzewski A., Krolikowski R., 1999, ”Noise Reduction in Audio Employing Auditory Masking Approach”, 106th Audio Engineering Society Convention, Munich, Germany, to be presented.

Johnston J., 1988, ”Transform Coding of Audio Signals Using Perceptual Noise Criteria”, IEEE Journal of Select. Areas in Communication, vol. 6, pp. 314-323.

Krolikowski R., Czyzewski A., 1999, ”Noise Reduction in Acoustic Signals Using the Perceptual Coding”, 137th Meeting Acoustic Society of America, Berlin, Germany, CD-Preprint.

Pawlak Z, 1982, ”Rough Sets”, International Journal of Computer and Information Sciences, No. 11, pp. 341-356.

Shlien S., 1994, ”Guide to MPEG-1 Audio Standard”, IEEE Trans. Broadcasting, vol. 40, pp. 206-218.
Vaseghi S., 1997, Advanced Signal Processing and Digital Noise Reduction, Wiley & Teubner, New York.

Zurada J., 1992, „Introduction to Artificial Neural Networks”, West Publishing Company, St. Paul.

Zwicker E., Fastl H., 1990, Psychoacoustics, Facts and Models, Springer Verlag, Berlin.

ACKNOWLEDGEMENTS

The research is sponsored by the State Committee for Scientific Research, Warsaw, Poland. Grant No. 8 T11D 021 12.