VYTAUTAS MAGNUS UNIVERSITY
VILNIUS UNIVERSITY
INSTITUTE OF MATEMATICS AND INFORMATICS

Gintarė ČEIDAITĖ

ACCOUSTIC SIGNALS RECOGNITION EQUIPMENT ADAPTATION TO THE VARIOUS ENVIRONMENTS

Summary of Doctoral Dissertation
Physical Sciences, Informatics (09P)

Kaunas, 2014

Dissertation was prepared at Vytautas Magnus University in 2007 – 2014.

Scientific supervisor:

prof. habil. dr. Laimutis Telksnys (Institute of Mathematics and Informatics, Physical Sciences, Informatics 09P).

The dissertation is to be defended in the Joint Council on Informatics of Vytautas Magnus University and Institute of Mathematics and Informatics.

Chairman:

prof. habil. dr. Vytautas Kaminskas (Vytautas Magnus University, Physical Sciences, Informatics 09P)

Members:

prof. habil. dr. Kazys Kazlauskas (Vilnius University, Physical Sciences, Informatics 09P),

prof. habil. dr. Feliksas Ivanauskas (Vilnius University, Physical Sciences,Informatics 09P),

prof. habil. dr. Dalius Navakauskas (Vilnius Gediminas Technical University, Technological Sciences, Informatics Engineering 07T),

prof. habil. dr. Rimantas Šeinauskas (Kaunas University of Technology, Technological Sciences, Informatics Engineering 07T)

Opponents:

prof. dr. Dalė Dzemydienė (Mykolas Romeris University, Physical Sciences, Informatics 09P),

prof. habil. dr. Rimvydas Simutis (Kaunas University of Technology, Technological Sciences, Informatics Engineering 07T).

The defence of dissertation is scheduled for 10 a.m. on September 26, 2014, at Vincas Čepinskis reading room, at the Faculty of Informatics of Vytautas Magnus University.

Address: Vileikos str. 8, LT-44404, Kaunas, Lithuania.

Summary of dissertation was mailed on August 25, 2014.

The dissertation is available at the National M. Mažvydas Library, the Library of Vytautas Magnus University, and the Library of Institute of Mathematics and Informatics.

VYTAUTO DIDŽIOJO UNIVERSITETAS
VILNIAUS UNIVERSITETO
MATEMATIKOS IR INFORMATIKOS INSTITUTAS

Gintarė ČEIDAITĖ

AKUSTINIŲ SIGNALŲ ATPAŽINIMO SISTEMŲ PRISITAIKYMO PRIE PASIKEITUSIŲ APLINKOS SĄLYGŲ TYRIMAS

Daktaro disertacijos santrauka
Fiziniai mokslai, Informatika (09P)

Kaunas, 2014

Disertacija rengta 2007 – 2014 metais Vytauto Didžiojo universitete.

Mokslinis vadovas:
prof. habil. dr. Laimutis Telksnys (Matematikos ir Informatikos Institutas, fiziniai mokslai, informatika 09P)

Disertacija ginama jungtinėje Vytauto Didžiojo universiteto ir Matematikos ir Informatikos instituto informatikos mokslo krypties taryboje.

Pirmininkas:
prof. habil. dr. Vytautas Kaminskas (Vytauto Didžiojo Universitetas, fiziniai mokslai, informatika 09P).

Nariai:
prof. habil. dr. Kazys Kazlauskas (Vilniaus universitetas, fiziniai mokslai, informatika 09P),
prof. habil. dr. Feliksas Ivanauskas (Vilniaus universitetas, fiziniai mokslai, informatika 09P),
prof. habil. dr. Dalius Navakauskas (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija 07T),
prof. habil. dr. Rimantas Šeinauskas (Kauno Technologijos universitetas, technologijos mokslai, informatikos inžinerija 07T).

Oponentai:
prof. dr. Dalė Dzemydienė (Mykolo Romerio universitetas, fiziniai mokslai, informatika 09P),
prof. habil. dr. Rimvydas Simutis (Kauno Technologijos Universitetas, technologijos mokslai, informatikos inžinerija 07T).

Disertacija bus ginama viešame gynimo tarybos posėdyje, kuris įvyks 2014 m. rugsėjo 26 d., 10 val. Vytauto Didžiojo universiteto Informatiko fakulteto Vinco Čepinskio skaitykloje.
Adresas: Vileikos g.8, LT-44404, Kaunas, Lietuva.
Disertacijos santrauka išsiuntinėta 2014 m. rugpjūčio 25 d.

Disertaciją galima peržiūrėti nacionalinėje M. Mažvydo, Vytauto Didžiojo universiteto, Matematikos ir informatikos instituto bibliotekose.

Notation

$\Omega_l(l=1,..,L)$	Dynamical system generated state
$X_i(l,p)$	Dynamical system signal of *l* state,genetated in the *p* environment
m	Order number of an autoregressive (AR) model
$x_i(l,p,\mu)$	μ -th sample of the generated signal $X_i(l,p)$
$A_p(p=1,2,...P)$	Environment
$a_j(l,p), b(l,p)$	parameters of the autoregressive model
$V_i(i=1,..,N)$	Gaussian random sequence; average $EV_i=0$, and dispersion $EV_i^2=1$
u	Dynamical system operation moment in the new environment
$\alpha_i(i=1,2,..)$	Parameter determining the moment of change (*u*) in the dynamical system environment
ζ_{S_1}, ζ_{S_2}	Variation boundaries of the parameter $\alpha_i(i=1,2,..,N)$
$E_{l_1}(l_1=1,...,L_1)$	Standards reflecting the states of dynamical systems
$e_t(l_1)$	Standard-forming *sub-standards* of state-reflecting indications
$D(j_1,i_1)$	Matrix describing indications of the standard, and the sample observed
$\delta(j_1,i_1)$	Weight of the distances of features
$\Delta(x_i(l,p,\mu),E_{l_1})$	Distance between the indications of sample $x_i(l,p,\mu)$, and the standard $E_{l_1}(l_1=1,...,L_1)$
$Zrc(x_i(l,p,\mu))$	Zero-crossing rate feature
$Se(x_i(l,p,\mu))$	Signal energy feature
c, d	Constants
dn	The amount of (sample) signal-warping frames
dl	Signal window frame length
$P(x_i(l,p,\mu))$	The quality of recognizing sample $x_i(l,p,\mu)$ in the *l-th environment*
$T_a(x_i(l,p,\mu))$	The amount of properly-recognized samples $x_i(l,p,\mu)$ in the *l*-th environment
r	The amount of samples under observation $x_i(l,p,\mu)$ in the *l*-th environment
d_s	The correct solution of the sample recognition
$F_{ii}(ii=1,..,\omega)$	Modeling of sequence for state sample
n_{11}, n_{12}	The amount of states $\Omega_l(l=1,..,L)$ for modeling sample emergence
ω	The general amount of states $\Omega_l(l=1,..,L)$ for modeling sample emergence
$\varphi(i_2=1,2....)$	Sequence of state-generated sample
ν	Sample sequence modeling multiplier
σ_p	Dispersion of the shortest distances $\alpha_p(p=1,..,P)$ obtained in $A_p(p=1,..,P)$ environment
$\overline{x_p}$	Average of the shortest distances $\alpha_p(p=1,..,P)$ obtained in $A_p(p=1,..,P)$ environment

INTRODUCTION

Relevance of the problem

The sphere of information technology (IT) is undergoing a phase of rapid development, thus, it offers ceaseless novelty, functional innovations and new possibilities. The main aim of the field nowadays is success in making IT more attractive as well as user-friendly, and the pace of the advancement being likely to retain its stability, the future is believed to bring mobile, easily-navigable and even voice-guided technological products.

Voice navigation is one of the main perspectives of further progress in the field. The current use of technology, such as laptops, tablets, smartphones, and other mobile devices is going to increase. People do already use and are likely to continue using their IT gadgets in cafes, bus stations, subways, underwater vehicles with particular atmosphere and in more or less noisy environment. And as voice recognition devices must successfully operate under different circumstances, it is very important to analyze the effects of different acoustic environments on acoustic signal recognition devices exploring the level to which the devices can adjust to the altering operation conditions successfully.

The dissertation analyzes the problem of acoustic signal recognition devices operation in non-stationary acoustic environments by providing the theory and constructive methods for ***evaluating acoustic attributions of enclosed environments***, which cause errors in sound signal recognition; and for ***precision in acoustic signal recognition maintenance*** when the acoustic signal recognition devices are found in environments of different acoustic attribution.

The impact of acoustic features of the environment on the quality of the operation of the acoustic signals recognition device is analyzed. The method of adjusting to altered circumstances and the implementing algorithms allowing the recognition of the moment of change in the environment features, and executing adjustment procedures under new conditions are explored.

In the dissertation, a theoretical model of the situation is presented. In it, the author indicates the possible ways of describing its constituent parts and reveals the way the acoustic signals are modeled. Also, the way of evaluating and modeling the acoustic specifications of the environment in which the recognition device operates is explained. More, the change in the qualities of speech signals found in the environment possessing new acoustic features is presented.

Software, based on the usage of the dynamic time warping method, was designed for conducting pilot researches. The software enables the modeling of various recognition situations in which both samples of the observed states, and the attributes of the environments in which the acoustic sound recognition processes take place can be controlled.

Conducting pilot researches on real speech signals is very expensive and it requires a great scale of human resources. Thus, the dissertation gives exceptional importance to acoustic signals that in a way reflect the attributes of speech signals. By modeling different possible recognition situations and states-generated samples, it is possible to examine diverse methods of adapting the teaching procedures. This allows to correctly predict when the model suggested is going to operate well, when it is difficult to adapt and under what circumstances the adaptation model suggested is not going to work properly. The answers to these questions are given in the dissertation in the form of pilot research results.

The effectiveness of the adaption to new environmental conditions method, based on the usage of recognition signals, is demonstrated in the pilot researches in the dissertation.

The aim of the work is to analyze the adaptation possibilities of the acoustic sound recognition device when adjusting to shifting environment conditions, and to propose a constructive methodology for adapting to the changed environment conditions.

In order to meet the aim of the work the following aspects are to be fulfilled:

1. To describe the acoustic attributes of the environments in which acoustic signal recognition processes (training and recognition) take place.
2. To analyze the impact of the acoustic characteristics of the environment on the recognition device.
3. To analyze the methods allowing to detect the moment of change in the characteristics of the environment.
4. To analyze the opportunities to operate training procedures of the recognition device in altered environments.
5. To create a pilot base with the help of which it would be possible to model various situations of acoustic signals recognition by controlling the acoustic characteristics of the environments.
6. To demonstrate the effectiveness of the proposed recognition device in adapting to the method of shifting environment conditions through pilot researches.

Scientific novelty

The following results attained in the dissertation are scientifically new:

1. The method of the acoustic signal recognition system's adaptation to the altered recognition conditions has been developed.
2. A software system allowing the examination of various acoustic signal recognition situations has been created.
3. Researches proving the effectiveness of the created method of acoustic signals' adaptation to changing environment have been carried out.

Practical significance

The methods proposed in the dissertation allow the functioning recognition system to adjust to the changing environment conditions retaining the quality of recognition.

The methodology of the work is based on the experimental research results. The method of adjusting to altering environment conditions puts an emphasis on the results provided by the recognition device, rather than on the particularities of operation of the device, thus, the method can be widely applicable to recognition devices using different recognition approaches.

It is possible to model various recognition situations, proceeding with the research work in the field of adjusting to shifting environment conditions, taking advantage of the software package developed in the dissertation.

Aprobation and publications of the research

The main results of the work were presented and discussed at the following conferences:

1. Analysis of Factors Influences Precision of Speech Recognition. 14th International Conference "Electronics 2010" May 18, 2010, Kaunas, Lithuania.

2. Estimation of the Environmental Impact on the Accuracy of Signal Recognition. The 19th International Conference on Information and Software Technologies (ICIST 2013), October 10th–11th, 2013, Kaunas, Lithuania.

Two publications are referred in scientific databases:

1. G. Čeidaitė. L. Telksnys Analysis of Factors Influencing Accuracy of Speech Recognition electronics and electronical engineering 2010. No. 9(105) 69 -72 psl. ISSN 1392 – 1215

2. G. Čeidaitė. L. Telksnys Estimation of the Environmental Impact on the Accuracy of Signal Recognition. Springer-Verlag CCIS (Communications in Computer and Information Science) Volume 403, 2013, 261-271 psl. ISBN 978-3-642-41946-1

Statements for the defence

The environment in which the acoustic signal recognition device operates has influence on the quality of recognition.

In order to ensure the recognition quality in the altered acoustic environments it is necessary to follow procedures of adjustment to new working conditions.

The adjustment to altered conditions process is influenced by the order of emergence of the states samples analyzed, and by the physical similarity of the states samples.

Using the recognition results given by the recognition device in order to control the quality of the recognition process is seen as appropriate.

The newly-developed methods (APMIA, NMPT, and Elipses) for adjusting to the alterred conditions open up possibilities to control the streams of altering recognition errors.

The experimental researches conducted prove the effectiveness of the proposed methods.

Structure and size of the work

The dissertation consists of an introduction, six chapters, conclusions, and a references list. The chapters are divided into subchapters some of which include several sections.

The introduction part of the dissertation describes the subject of the research and relevance of the topic. The part also involves the formulation of the aim of the work, description of research methods, scientific novelty, practical significance of the dissertation, approval of the work, a list of research publications, an overview of the structure of the work, and the content of the dissertation.

The first chapter gives an overview of the works of the scientists of both Lithuanian and foreign origin conducted in the thematically-related sphere. The works carried out in the field of acoustic signals recognition while examining the impact of the existing noises on the quality of recognition and possibilities of adjustment to them are analyzed in the dissertation.

The second chapter examines factors exerting influence over the acoustics of the environments. Also, the method for measuring and describing acoustic environment characteristics is presented. More, the method for modeling acoustic signals in environments possessing particular acoustic characteristics is discussed. An experimental research the results of which demonstrate the impact of the environment where the device operates on the acoustic signal recognition device is described.

The third chapter gives an analysis of the problem of establishing the moment of change in the environment conditions. A parameter the observation of which allows the establishment of the moment of change in the environment conditions is presented. Methods for generating the observed state samples in the environment are described. Also, a methodology for evaluating environment conditions is provided.

The fourth chapter includes the discussion of three constructive methods with the help of which the recognition device can easily adapt to and effectively operate in the new environment. A detailed overview of each respective method's actions the fulfillment of which allows the adjustment of the recognition device to the new situation is proposed.

The fifth chapter describes the software with the help of which the characteristics of the observed states samples, as well as their sequence of emergence in the environment during the recognition process is possible to model. With the help of the software, various recognition situations and ways of adjusting to new situations can be modeled. This particular software has been used while conducting experimental researches demonstrating the effectiveness of the method while adjusting to changeable conditions.

The sixth chapter describes experimental researches that have been conducted. The chapter also provides a detailed description of the state samples used and the process of modeling them in environments possessing new acoustic characteristics. There was an experimental research conducted seeking to establish which of the processes given in the third chapter should be used in order to identify the moment of change in the characteristics of the environment within the shortest time possible. Results and conslusions deriving from the research are provided. Also, an experimental research analyzing the effectiveness of adjusting to the altering environment conditions is described. The methods already discussed in the fourth chapter are analyzed questioning their effectiveness; results and conclusions of the research are presented.

CONTENT OF THE DISSERTATION

1 Evaluating acoustic characteristics of environments

There are numerous factors affecting the sounds traveling within the environment. Thus, in this particular situation, it is worthwhile to describe the acoustic characteristics of environments by giving an overview of the acoustic signals' (sounds') amplitude spectral characteristics of the enviroment

$$S(f)=\sqrt{\left(\sum_{k=-\infty}^{\infty}h(k)\cos 2\pi fk\right)^2-\left(\sum_{k=-\infty}^{\infty}h(k)\sin 2\pi fk\right)^2}\,, \tag{1.1}$$

Where $h(k)$ – acoustic response signal function of time describing the acoustic characteristics of environments. Acoustic response signal function $h(k)$ defines the response $y(n)$ by the signal $x(n)$ within the environment:

$$y(n)=\sum_{k=0}^{\infty}h(k)x(n-k) \tag{1.2}$$

If the signal $x(n-k)$ is of short impulse $\delta(n-k)$ the function of which is

$$x(n-k)=\delta(n-k)=\begin{cases}1, & n=k\\ 0, & n\neq k\end{cases}, \tag{1.3}$$

then

$$y(n)=\sum_{k=0}^{\infty}h(k)\delta(n-k)=h(k)\,. \tag{1.4}$$

The acoustic response signal function $h(k)$ might be acquired by releasing an impulse signal $\delta(n)$ in the environment. Impulse signal $\delta(n)$ [15] has been used during the experiments.

$$\delta(n)=4\pi\left(\frac{m}{2\pi K_B T}\right)^{\frac{3}{2}}n^2 e^{\frac{mV^2}{2K_B T}} \tag{1.5}$$

Here K_B – the *Boltzmann constant*, m – gass molecular mass, T – temperature, V – gas molecular speed. The impulse sinal has been generated by shooting a 9 mm revolver depicted in Figure 1.1.

Figure 1.1 A 9 mm revolver used for analyzing the acoustic characteristics of the environments

The spectral characteristics $S_A(f), S_B(f), S_C(f), S_D(f)$ of environments A, B, C and D are presented in Figure 1.2:

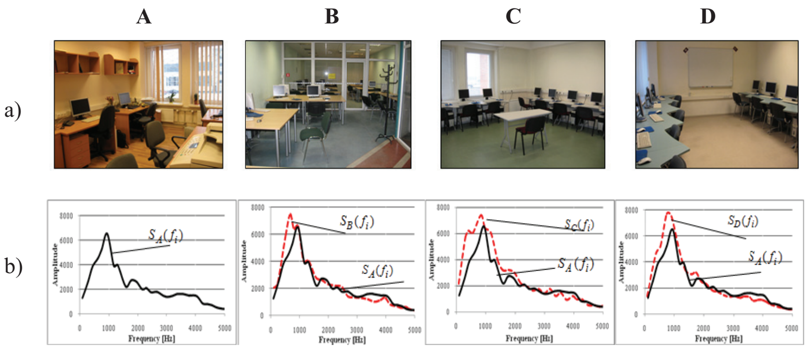

Figure 1.2 a) The environments of the experimental research b) Spectral characteristic $S(f_i)$ of environments A, B,C, D.

2 Research on the environment's impact on acoustic signals

Acoustic signals recognition devices are trained in one particular environment, thus, happening to operate in a different acoustic environment, their accuracy of the devices' operation might be inadequate due to the impact of the new environment's specific characteristics. This is illustrated by the below-given example.

A problem in which we deal with the acoustic signal recognition device trained to recognize acoustic speech signals in the environment A is to be solved (Figure 1.2). During the time moment *u*, the recognition device moves from the Environment **A** to Environments **B, C** and **D**. We are to analyze the impact the change in the environments has on the accuracy of recognition.

Training the recognition system

Two recognition situations were analyzed in the experiment. To train the class a set of 10 speech signal examples was used. Pattern references signals created in the Environment A by woman voice at the same time, and different size sets of speech signal pattern references were used for all training situations

1st training situation. Two first speech signal pattern references were taken from the set of speech signals for training. Recognizer used the nearest neighbor method to choose the best signal for training. In this situation we had a recognizer trained with one Ω class pattern reference of speech command.

2nd training situation. Six first speech signal pattern references were taken from a set of speech signals for training. Recognizer used the same nearest neighbor method to choose five best signals for training. In this situation we had a recognizer trained with five Ω class pattern references of speech command.

The basic environment for training was Environment A (signed EA). One female utterance was used for training. Recognition process was performed in environments A, B, C and D.

Solving the recognition problem

A signal recognition system, identifying separately-uttered speech signals which is based on the dynamic time scale warping method, was used in the research [23]. Its operation constitutes three parts: speech signal retrieval, training the class according to the references and speech signal identification.

Once the speech signal of the speaker gets into the recognition system, it is processed. At first, initial and final sound boundaries are identified. This is done using the method of dynamic programming. By using Bellman functions, the most probable moments of signal change are found – these are then taken as initial and final boundaries of the speech signal. The rest part of the speech signal is regarded as a signal of the environment.

The training of the recognizer is based on the closest neighboring sound principle. From a variety of training examples, a set of examples with smallest distances between them are selected. Then, the calculation is done in the following manner: firstly, calculations with all combinations of the possible reference are done; one of the training utterances is taken as a reference, while the rest are regarded as comparatives. Then, it is established, which of the training examples shows the best results, i. e. which example has the shortest average distance to other examples. Afterwards, calculations with twos of possible references are done – the training examples two showing the best results is ascertained. The same procedure is carried out with the training examples threes and fours.

The dynamic time scale warping method is used while establishing the shortest path to the class reference. Itakura rule is used when examining the path of the signal [23].

Objects of the study

There are 80 speakers of different contingent participating in the research, their ages varying from 20 to 55 years. Two speech signals states $\Omega_l (l = Kaunas, Vilnius)$

(we denote it as $\Omega_l(l = K,V)$) are analyzed in the research. Every speaker produces approximately 25 speech signal examples, so all in all there are $\Omega_l(l = K) = 2910$ and $\Omega_l(l = V) = 2945$ examples. The signals are recorded at the sampling rate 11025 Hz, using 16 bit mono format.

Results

During the research, two speech signals states $\Omega_l(l = K,V)$, 4 different environments and 80 speakers of different age are analysed. Three types of features were used [23]:

- cepstrum coefficients,
- cepstrum coefficients with mean subtraction,
- autoregressive model coefficients

Two reference training situations are analyzed: in the first one, the states $\Omega_l(l = K,V)$ are trained according to the example of a single reference $E_{l1}(l_1 = 1,2)$ whereas in the second, the states are trained with examples of five references $E_{l1}(l_1 = 1,2)$. Two recognition situations are present. The first one takes place when both training and recognition of the references $E_{l1}(l_1 = 1,2)$ takes place in one particular environment and in the second situation, training and recognition of the references $E_{l1}(l_1 = 1,2)$ takes place in different environments. Figure 2.1 a, b, c illustrates the results of the experiment.

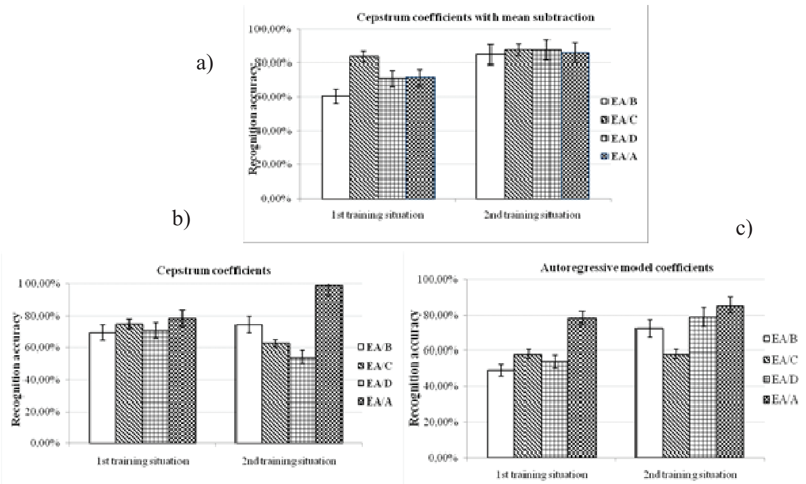

Figure 2.1 Recognition results using a) cepstrum coefficients with mean subtraction features b) cepstrum coefficients features c) autoregressive model coefficients

Environment A is a reference environment (we denote it as EA); in it, female speech signals are recorded in order to be used while training the recognition system. The acoustic speech signals recognition process takes place in the following enviroonments: A, B, C and D. The environments used in the process of the speech signals system training process are taken as reference environments, and in the following work these are indicated as EA, EB, EC, and ED. Figure 2.11 depicts research results which demonstrate a decline in the recognition quality as training and assessment environments change.

Conclusion

Having carried out an analysis of environment acoustic characteristics, it was clarified that the acoustics of environments has a great effect on the acoustic signals recognition. The impact cannot be minimized or avoided even by adapting various recognized signal characteristics systems. The results of the research have shown that the recognition quality declines by 20% when training and recognition take place in different environments. And yet, although it is impossible to avoid the decline in recognition quality, slightly better recognition results might be acquired while using systems of centered cepstrum coefficient characteristics. Thus, it is important to investigate the potential of acoustic signals recognition devices to adapt to changing environment conditions.

3 Detecting acoustic changes in recognition environments

The chapter analyzes one of the main problems of the dissertation – determining the time momentum u identifying the change in environment characteristics. The recognition device is trained to recognize in the environment $A_p(p=1)$ functioning the samples $x_i(l,p,\mu)$ generated by the dynamic system. The samples belong to states $\Omega_l(l=1,2,..L)$, in which $l=1,2,..L$ describes a state, $p=1,2,..P$ indicates the environment in which the sample $x_i(l,p,\mu)$ was generated by the dynamic system – μ – is the rank number of sample $\mu=1,2,.....$. During the time moment u the dynamic system helps to function in the environment $A_p(p=2)$.

The recognition device, once it recognizes the input sample $x_i(l,p,\mu)$, returns the difference - the estimate of the parameter $\alpha_i(i=1,.2,...)$- between the two shortest distances in the environment $A_p(p=1)$ and $A_p(p=2)$ respectively. Depending on this, the time moment $i=u$ has to be detected. It indicates the recognision device starting to recognize in the new environment $A_p(p=2)$, generating samples $x_i(l,p,\mu)$ which belong to the states $\Omega_l(l=1,2,..L)$ in which $l=1,2,.L$ describes the state, $p=1,2,..P$ – the environment in which the dynamic system has generated sample $x_i(l,p,\mu)$, μ being the rank number $\mu=1,2,...$ of this sequence. The situation under discussion is illustrated in Figure 3.1.

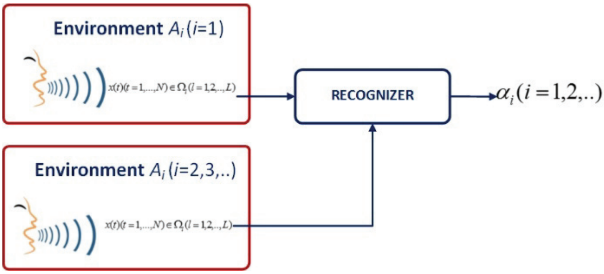

Figure 3.1 The scheme of the situation analyzed

Evaluating the parameter α_i identifying acoustic environment characteristics

While carrying out experimental acoustic signal recognition researches and looking for regularities which would help to evaluate and characterize signals occurring in a particular environment, an observation has been made that when the recognition device, trained to recognize particular signals, is being placed in a specific environment, where particular conditions prevail, the difference values between the signal observed and reference characteristics turn out to be very similar. So, when observing objects in a particular environment, the shortest distances difference values possible to obtain might be predicted. This regularity might be used as a critical indication while determining the moment u of the systems changed environment characteristics.

Paremeter α_i is calculated according to the following algorithm:

Sample $x_i(l, p, \mu) \in \Omega_l(l = 1,2,..L)$ is generated by the dynamic system. In it, $l = 1,2,..L$ describes the state, $p = 1,2,..P$ is the environment in which the dynamic system has generated the sample $x_i(l, p, \mu)$, μ being the rank number of the sample $\mu = 1,2,.....$ getting into the recognition device which is trained to recognize samples of this kind.

The recognition device is trained to recognize samples $x_i(l, p, \mu)$ belonging to the state $\Omega_l(l = 1,2,..L)$. The training is accomplished by providing the training procedure with reference examples $E_l(l_1 = 1,...,L_1)$ reflecting states $\Omega_l(l = 1,..., L)$.

$$E_{l_1}(l_1 = 1,...L_1) = \{e_1(l_1), e_2(l_1),...., e_h(l_1)\} \tag{3.1}$$

State references $E_{l_1}(l_1 = 1,...L_1)$ are compiled from the sub-references $e_t(l_1)$ formed by the features set.

$$e_t(l_1) = \{Zcr(x_z(l, p, \mu), Se(x_z(l, p, \mu)\} \tag{3.2}$$

, in which $k = 1,.., dn$, $t = 1,...,h$, h being the amount of reference $E_{l_1}(l_1 = 1,...L_1)$ forming sub-patterns. $Zcr(x_z(l, p, \mu))$ is the zero-crossing rate characteristic system, it is calculated as follows:

$$Zcr(x_z(l,p,\mu)) = \sum_{k=1}^{dn} \frac{\left|\operatorname{sgn}(x_z^k(l,p,\mu)) - \operatorname{sgn}(x_z^{k-1}(l,p,\mu))\right|}{2} \tag{3.3}$$

$$Sgn(x_z^k(l,p,\mu)) = \begin{cases} 1, & x_z^k(l,p,\mu) \geq 0 \\ -1, & x_z^k(l,p,\mu) < 0 \end{cases} \tag{3.4}$$

$Se(x_z(l,p,\mu))$ is the system of signals energy characteristics which is calculated as follows:

$$Se(x_z(l,p,\mu)) = \sum_{k=1}^{dn} (x_z^k(l,p,\mu))^2 \tag{3.5}$$

During the process of indications formation, signal $z = 1,...,Z$ from the state $\Omega_l(l = 1,...,L)$ sample $x_z(l,p,\mu)$ is divided into *dn* parts of *dl* signal length. $l = 1,2,..L$ describes the state, $p = 1,2,..P$ environment, in which the dynamic system generates sample $x_i(l,p,\mu)$, μ being the rank number $\mu = 1,2,.....$ of this sample.

Sub-reference $e_t(l_1)$ characteristics set-forming features are discerned within each part $x_z^k(l,p,\mu)$, $k = 1,..,dn$ of sample signal $x_z(l,p,\mu)$ by using feature systems.

The feature systems used do not require the recognition system to possess high scale resources, thus the systems are useful while computing artificial speech signals. Naturally, it is also possible to use other indications.

The above-discussed states-describing method of reference selection is based on an assumption that the recognition device is given precise state-defining sample examples which are taken as reference examples. It is also possible, however, to use a different method of reference selection - the recognition device is given a few state describing sample examples by using the recursive reference selection method with the help of which, selection of the least similar state-descriptive reference examples is enabled. The procedure includes comparing every reference example with every already-enlisted one which leads to making a decision about which of the references is the least similar one, i.e. which demonstrates the biggest difference when comparing state-descriptive samples. The following is the reference list selection algorithm:

Particular state $\Omega_l(l = 1,...,L)$ reflecting references $E_l(l = 1,...,L)$, where $E_l(l = 1,...) = \{e_1(l), e_2(l),....,e_h(l)\}$, are observed in order to select the least similar state references. Every reference's sub-reference has to be compared to the every enlisted reference. The dynamic time warping method is used for making the comparison. Zero-crossing rate and short-time energy indications are going to be used.

Initialization:

$$e_i'(l) = E_l(1)$$

Recursion:

$$e\max(i) = \max(E_l(l = 1,...), e_i'(l))$$

$$, kai \quad i = 1,2,...N$$

Termination:

$$E_l(l = 1,...) = \arg\max(e\max(i = 1, 2,..), h)$$

, when *h*- is the amount of the selected references and i – is the rank number if reference comparison iteration.

With the help of this algorithm, as diverse set of particular state reflecting references as possible is formed during the training process. With the help of the primal algorithm, provision of only the same or very similar reference examples to the training system is possible, while by using this particular reference selection method, reference set can be diversified by adding different examples. In this manner the quality of recognition in the recognition process is improved.

Figure 3.2 illustrates a scheme of the input sample $x_i(l,p,\mu)$ for measuring the distance $\Delta(x_i(l,p,\mu),e_t(l_1))$, where $l=1,2,..L$ describes the state, $p=1,2,..P$ describe the environment in which the dynamic system has generated sample $x_i(l,p,\mu)$ – μ being the rank number $\mu=1,2,.....$ of this sample, and l_1-reference describing state being $\Omega_l(l=1,2,...,L)$.

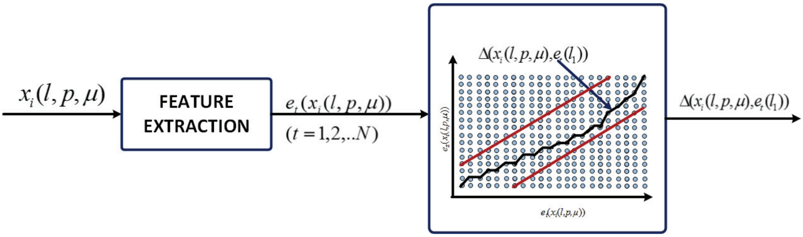

Figure 3.2 Scheme for calculating distance $\Delta(x_i(l,p,\mu),e_t(l_1))$ in the sample $x_i(l,p,\mu)$

In order to detect distance $\Delta(x_i(l,p,\mu),E_{l_1})$ of the sample $x_i(l,p,\mu)$ to the references' E_{l_1} describes the dynamic system's states, the Sakoe-Chiba method [26] in which $\Delta(x_i(l,p,\mu),E_{l_1})$ is calculated as ($l=1,2,..L$ state, $p=1,2,..P$ environment, and μ rank number of this sample) :

1. Feature comparison matrix constituent distance weights $\delta(i_1,j_1)$, in which element i_1 describes the *k*-th features of short-time energy $Se(x_i^{i_1}(l,p,\mu))$, $i_1=1,2,..dn$ and j_1 indicates the attributes $j_1=1,2,..dn$ of *k*-th section zero-crossing rate $Zcr(x_z^{j_1}(l,p,\mu))$ in the input sample $x_i(l,p,\mu)$, are evaluated. Distance weights $\delta(i_1,j_1)$ are calculated as follows:

$$\delta(j_1,i_1)=\sqrt{Se(x_z^{j_1}(l,p,\mu))-Se(x_i^{i_1}(l,p,\mu)))^2-(Zcr(x_z^{j_1}(l,p,\mu))-Zcr(x^{i_1}(l,p,\mu)))^2} \quad (3.6)$$

All feature distances between every sub-reference $e_t(l_1)$ of the observed sample $x_i(l,p,\mu)$ are evaluated

$$\Delta(x_i(l,p,\mu),e_t(l_1)) = \min([D(j_1-1,i_1-1)D(j_1,i_1-1)D(j_1-1,i_1)] + \delta(j_1,i_1)) \quad (3.7)$$

2. Feature distance $\Delta(x_i(l,p,\mu),E_{l_1})$ to the dynamic system state-reflecting references $E_{l_1}(l_1=1,..,L_1)$ of the input sample $x_i(l,p,\mu)$ is calculated

$$\Delta(x_i(l,p,\mu),E_{l_1}) = \min_{t=1,..,h} \Delta(x_i(l,p,\mu),e_t(l_1)) \quad (3.8)$$

The acquired shortest distances' $\Delta(x_i(l,p,\mu),E_{l_1})$ estimates are sorted in the advancing order (Figure 3.3) selecting two least estimates β_1 and β_2 of the distance $\Delta(x_i(l,p,\mu),E_l)$.

$$\beta_1 = \min_{t=1,...,h} \Delta(x_i(l,p,\mu),E_l) \quad (3.9)$$

$$\beta_2 = \min_{t=1,...,h} (\Delta(x_i(l,p,\mu),E_l) - \beta_1) \quad (3.10)$$

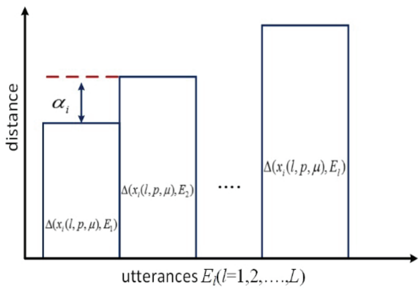

Figure 3.3 Scheme for evaluating parameter α_i

Having two shortest distance estimates β_1 and β_2 allows calculating parameter α_i which is the difference between these two estimates:

$$\alpha_i = \beta_2 - \beta_1 \ , i=1,2,... \quad (3.11)$$

4 Research on recognizers' adaptation to the variable environmental conditions

Estimate sequence of the parameter $\alpha_i (i=1,2,...)$ (3.11) given by the recognition device is observed; on the basis of it, the time momentum $i=u$ indicating the starting point of the recognizer operation in the new environment $A_p(p=2..,P)$. Here $l=1,2,..L$ describes the state, $p=1,2,..P$ describe the environment in which the dynamic system has generated sample $x_i(l,p,\mu)$, μ being rank number $\mu = 1,2,.....$ of this sample (Figure 4.1).

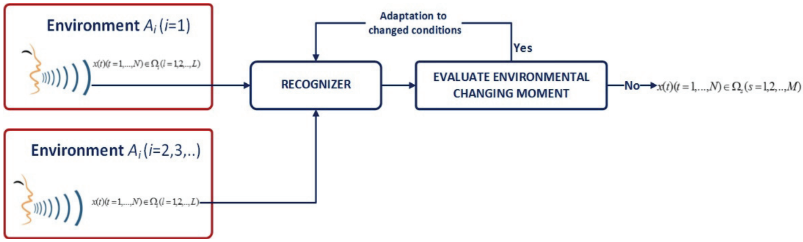

Figure 4.1 Scheme of the situation analyzed

Having detect the moment $i = u$ of change in the environment $A_p(p = 1,2,..,P)$ conditions, the recognition device has to adapt to recognizing state $\Omega_l(l = 1,2,..L)$ samples $x_i(l,p,\mu)$ generated by the dynamic system functioning in a new environment $A_p(p = 2,..,P)$. This chapter is going to analyze the methods of the acoustic signal recognition device adapting to changing conditions. According to the aforementioned, once the time momentum $i = u$ of change in environment conditions is set, the recognizer can adapt to newly-created conditions and continue recognition processes.

4.1 *Method for identifying the moment of change in features (APMIA)*

The method is compiled of two parts. ***Part one -*** identifying the moment u of change in the environment conditions and adapting the device's training procedure to the changed conditions. ***Part two -*** refining the recognition device's training procedure by adding new state sample examples that have emerged in the environment.

Experimental researches have shown that training the recognition device to recognize one signal of particular state is not a condition of adequate training. Thus, the recognizer has to be exposed to several additional state samples. Given below is the modelled algorithm basing on which enables us to solve the analyzed problem of adapting to changing conditions.

1. Initial recognition conditions for the recognizer are formed:
   1.1 Recognition device is trained to identify state elements $x_i(l,p,\mu)$ generated by the dynamic system. In the eqution, $l = 1,2,..L$ describes the state, $p = 1,2,..P$ states the environment in which the dynamic system has generated sample $x_i(l,p,\mu)$, μ being the rank number $\mu = 1,2,.....$ of this sample.
   1.2 Features identifying environment $A_p(p = 1,2,...P)$ under discussion are defined.
2. State element $x_i(l,p,\mu)$ generated by the dynamic system and that needs to be identified is given to the recognition device.
3. Recognition device operates on the following:

3.1. State $\Omega_l(l=1,2,..L)$ belonging to the observed sample $x_i(l,p,\mu)$ is determined;

3.2. According to the present environment $A_p(p=1,2,..P)$ identifying data, it is estimated whether the observed state sample $x_i(l,p,\mu)$ has been generated in the same environment $A_p(p=1,2,..P)$.

3.2.1. If the environment $A_p(p=1,2,..P)$ characteristics' change momentum u is not estimated, decision on recognition is made (the procedure goes to point 4).

3.2.2. If the environment change momentum u is estimated, then the procedure of adapting to new conditions follows (fulfilling point 5).

4. Recognition result is drawn. The signal is assigned to a particular state.
5. Adaptation to the changed conditions procedure is carried out.

5.1. The impact of the environment $A_p(p=2,..P)$ over the observed sample $x_i(l,p,\mu)$ is subdued.

5.2. Features $A_p(p=1)$ of the primal environment condition are assigned to the acquired signal (procedure of signals'summation is carried out, i.e., the environment is added up to the observed signal).

5.3. The acquired signal $x_i(l,p=1,\mu)$ is returned to the recognition device for establishing its state.

5.4. Having identified the state l of the signal state, the procedure of recognizer's adaptation to new conditions is carried out:

5.4.1. State $\Omega_l(l=1,2,...,L)$ describing reference list $E_l(l_1=1,...,L_1)$ is specified.

5.4.2. Environment $A_p(p=2,..P)$ identifying parameters are specified. I.e., the parameter's α_i change boundaries ξ_{S_1} and ξ_{S_2} in the environment $A_p(p=2,..P)$.

6. Once new dynamic system generated samples $x_i(l,p,\mu)$ $x_i(l,p,\mu)$ emerge and are identified, procedures of specifying the training process of the recognition device are carried out without establishing the moment u of change in environment conditions.

6.1. The quality of the recognition device identifying states $\Omega_l(l=1,2,...,L)$ is observed.

6.1.1. If the worsening of the recognition quality is detected, the procedure of state samples list comparison is carried out - comparison of state samples list of the reference signals defining particular dynamic system's state element is conducted. The least state-resembling element of the list is eliminated.

6.1.2. In case the worsening in the quality of the work of the recognition device is not detected, the work is proceded.

6.2. Reference list is updated specifying it with several examples of signal samples in the new environment $A_p(p=2,..P)$.

6.2.1. The change boundaries ξ_{S_1} and ξ_{S_2} of the parameter α_i are examined:

6.2.1.1. If the criterion is satisfied, reference list is specified.

6.2.1.2.

6.2.1.3. If the criterion is dissatisfied, specification of the reference list is not conducted.

6.2.2. The rank number of the samples list μ of the signal $x_i(l, p, \mu)$ observed is evaluated. In case it exceeds the previously-set value, for instance, 30 and more, the boundaries ξ_{S_1} and ξ_{S_2} of the environment identifying parameter α_i are recalculated.

6.3. Decision over recognition is made.

Below given (Figure 4.2) is the scheme of the previously-discussed algorithm

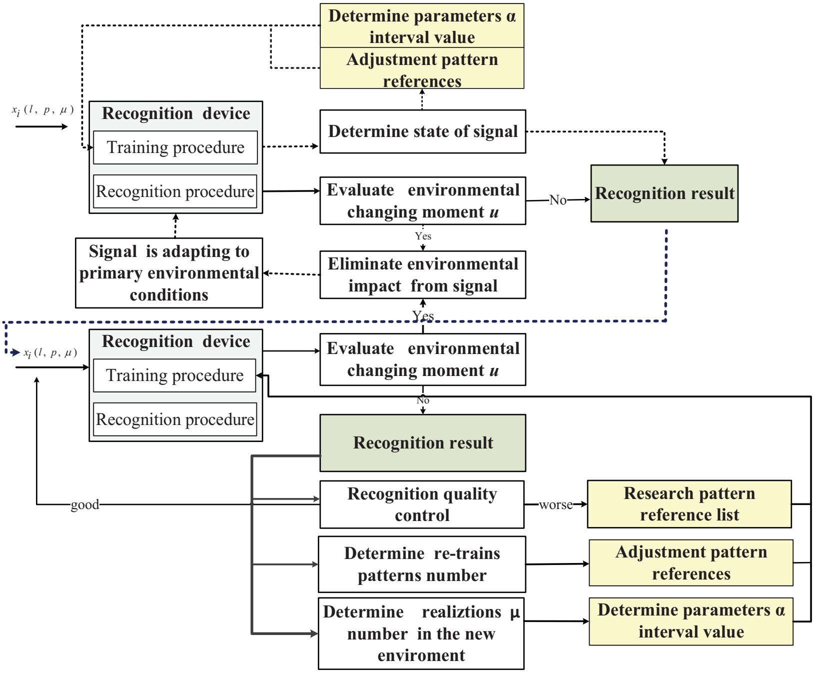

Figure 4.2 A model of the adaptive signal recognition system

Individual constituent parts of this model adaptive to changing recognition conditions are going to be discussed further in the chapter.

4.2 *Constant specification of the training system (NMPT)*

Acoustic signals of speech possess very distinctive qualities. Thus, it is very difficult to use artificial sound-forming methods, as none of the artificially-created signals is likely to reflect the real signal characteristics. Several additional factors must be paid attention to willing to retain the quality of recognition, making decisions and building on the results given by the recognition device. It is recommended to store all the correctly-recognized signals in the additional "signal storage" information of which would be adaptable to state reference list specification. Below given is a scheme illustrating the way a recognition system conducts the recognition of original acoustic speech signals.

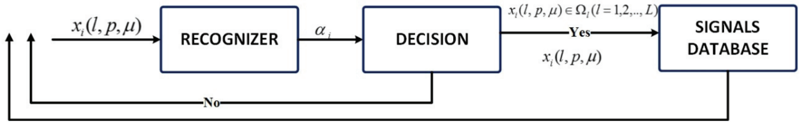

Figure 4.3 A model of correctly-recognized signals

A scheme presents the model when a signal $x_i(l,p,\mu)$ in which $l=1,2,..L$ describes the state, $p=1,2,..P$ - environment, and μ is the rank number, belonging to the particular state $\Omega_l(l=1,2,...,L)$, gets into the recognition device. The device, basing on the examples of state references from the training procedure and on the procedure itself, returns the solution of the recognition problem.

Recognition device can make solutions of three kinds:

- Recognition solution has been accepted,
- Recognition solution has been declined,
- Recognition solution cannot be accepted due to a variety of possible reasons; when the result is dubious, it is discarded trying to repeat the recognition procedure with the state signal of a subsequent sample.

If the recognition result is accepted, the signal recognized is referred to as an example belonging to a particular state this enabling it to be transferred into the signals storage database where it will afterwards be used as an example during the training procedure. If the recognition result is declined, the recognition procedures are continued.

This algorithm helps the recognition device to adapt to a particular situation, but there is always a possibility that during every review the necessary state $\Omega_l(l=1,2,...,L)$ defining sample $x_z(l,p,\mu)$ might be eliminated from the reference list.

4.3 Method for identifying the moment of change in features and *check dependence ellipse area(* Ellipses)

The method of Ellipse adaptation operates similarly to the method of evaluating the moment *u* of change in conditions of the environment; the difference between the methods being the following: before making the recognition problem solution $d_s: x_i(l,p,\mu)\in\Omega_l(l=1,2,..,L)$, the shortest distance α_i to reference samples $x_z(l,p,\mu)$ of states $\Omega_l(l=1)$ and $\Omega_l(l=2)$ is measured checking wheter function $f_{elip}(l,p,\mu)$ incorporates it.

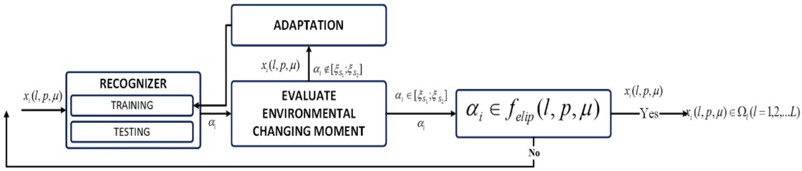

Figure 4.4 A diagram of Ellipse method adaptation to the changing conditions

In this case Ellipse serves as a classificatory with the help of which it is possible to evaluate whether the result given by the recognition device is correct, or not. Thus, the quality of the recognition process might be improved by evaluating the ellipse boundaries belonging to a particular state of the initial recognition moment, and by drawing the state identification result owing to these boundaries during the recognition process. It is also not complicated to adapt ellipse function in the new environment – it is possible to recalculate it if the data is present.

It is recommended to calculate function of a new ellipse by carrying out transformations on turning point and impulse (Figure 4.6).

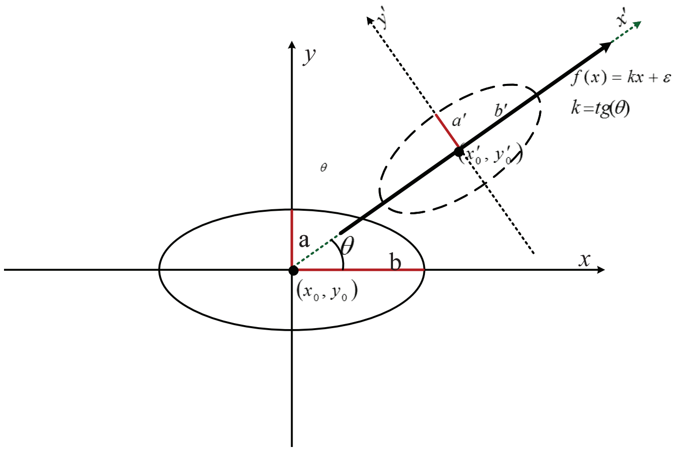

Figure 4.6 A model of Ellipse transformation

Since the pairs of observation are present, i.e. we have the shortest distances to states $\Omega_l(l=1)$., and $\Omega_l(l=2)$, regression equation can be used for evaluating data dependence [43].

$$f(x) = kx + \varepsilon \qquad (4.1)$$

, in which k - straight's direction coefficient, θ - describes the size of direction angle, ε - the segment intercepted by the straight in the **Oy** axis [43].

$$k = tg(\theta) \tag{4.2}$$

The foci coordinates (x_0', y_0') of the coordinates' system of the ellipse transformed are calculated [43]

$$x_0' = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{4.3}$$

$$y_0' = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{4.4}$$

, in which x_i are the distance values situated in the X-axis, and y_i are the distance values situated in the Y-axis.
Having evaluated dispersions σ_x, σ_y of the data sample, with the help of which the breadth of data distribution can be measured, ellipse's big and small half axles a', b' [43] are calculated.

$$a' = \frac{\sigma_y}{2} \tag{4.5}$$

$$b' = \frac{\sigma_x}{2} \tag{4.6}$$

Ellipse equation is formulated:

$$\frac{x'^2}{a'^2} + \frac{y'^2}{b'^2} = 1 \tag{4.7}$$

, where the foci F_1', F_2' of ellipse are externalized

$$\begin{gathered} a'^2 = b'^2 + c'^2, \\ F_1'(-c', 0), \quad F_2'(c', 0) \end{gathered} \tag{4.8}$$

4.4 *Measures for evaluating algorithm effectiveness*

Seeking to evaluate the effectiveness of the algorithm of the recognition device analyzed to adapt to changed environment conditions, measures for evaluating the quality are determined:

-The speed of adapting to the changed conditions
-Recognition quality in the new environment
-Evaluating the moment of change in the environment
-The amount of the recognizer's recognition errors

4.5 *Reasons for the lowering recognition quality*

Several factors can influence the lowering in the recognition quality of the recognition device after adaptation to new environment conditions:

- The moment u of change in the environment conditions was not determined in due course. Consequently, the recognition device does not know that the adaptation to new conditions has to be started.

- While adapting, the reference list is specified training the system according to examples belonging to another state. A solution to the problem could be a constant revision of the reference signals list.
- Selecting the boundaries ξ_{S_1} and ξ_{S_2} defining environment features - in case the boundaries are selected incorrectly, it is very difficult to evaluate the moment u of changed environment conditions. Faulty reference list training is also possible.

5 Research on recognizers' adaptation

5.1 *Modeling state* $\Omega_l(l=1,2)$ *generated samples in environments* $A_p(p=1,2)$

In the research state $\Omega_l(l=Kl4)$ samples $x_i(l,p,\mu)$ signals with $\Omega_l(l=Kl1,Kl2,Kl3,Kl5,Kl6,Kl7)$ in respect to samples $x_i(l,p,\mu)$ are going to be analyzed. State $\Omega_l(l=Kl1,Kl2,Kl3,Kl5,Kl6,Kl7)$ samples $x_i(l,p,\mu)$ functioning in the environment $A_p(p=1)$ are defined in equations:
State $\Omega_l(l=Kl1)$ sample $x_i(l,p,\mu)$ in the environment $A_p(p=1)$ is defined**:**

$$x_i(1,1,\mu) = -1.19x_{i-1}(1,1,\mu) + 0.74x_{i-2}(1,1,\mu) + v_i(i=1,...,n) \quad (5.1)$$

, its spectral density function $S_{Kl5}(f_i)$ is defined:

$$S_{Kl1}(f_i) = \frac{1}{11025}\frac{2}{\left|1+1.19^{\frac{-j2\pi f_i}{11025}} - 0.74e^{\frac{-j4\pi f_i}{11025}}\right|^2} \quad (5.2)$$

State $\Omega_l(l=Kl2)$ sample $x_i(l,p,\mu)$ in the environment $A_p(p=1)$ is defined**:**

$$x_i(2,1,\mu) = -1.18x_{i-1}(2,1,\mu) + 0.85x_{i-2}(2,1,\mu) + v_i(i=1,...,n) \quad (5.3)$$

, its spectral density function $S_{Kl2}(f_i)$ is defined:

$$S_{Kl2}(f_i) = \frac{1}{11025} \frac{2}{\left|1 + 1.18^{\frac{-j2\pi f_i}{11025}} - 0.85e^{\frac{-j4\pi f_i}{11025}}\right|^2} \tag{5.4}$$

State $\Omega_l(l = Kl3)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p=1)$ is defined:

$$x_i(3,1,\mu) = -0.143x_{i-1}(3,1,\mu) + 0.801x_{i-2}(3,1,\mu) + v_i(i = 1,...,n) \tag{5.5}$$

, its spectral density function $S_{Kl3}(f_i)$ is defined

$$S_{Kl3}(f_i) = \frac{1}{11025} \frac{2}{\left|1 + 1.143^{\frac{-j2\pi f_i}{11025}} - 0.801e^{\frac{-j4\pi f_i}{11025}}\right|^2} \tag{5.6}$$

State $\Omega_l(l = Kl4)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p=1)$ is defined:

$$x_i(4,1,\mu) = -1.24x_{i-1}(41,\mu) + 0.831x_{i-2}(4,1,\mu) + v_i(i = 1,...,n) \tag{5.7}$$

, its spectral density function $S_{Kl4}(f_i)$ is defined

$$S_{Kl4}(f_i) = \frac{1}{11025} \frac{2}{\left|1 + 1.24^{\frac{-j2\pi f_i}{11025}} - 0.831e^{\frac{-j4\pi f_i}{11025}}\right|^2} \tag{5.8}$$

State $\Omega_l(l = Kl5)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p=1)$ is defined:

$$x_i(5,1,\mu) = -0.957x_{i-1}(5,1,\mu) + 0.767x_{i-2}(5,1,\mu) + v_i(i = 1,...,n) \tag{5.9}$$

, its spectral density function $S_{Kl5}(f_i)$ is defined

$$S_{Kl5}(f_i) = \frac{1}{11025} \frac{2}{\left|1 + 0.957^{\frac{-j2\pi f_i}{11025}} - 0.767e^{\frac{-j4\pi f_i}{11025}}\right|^2} \tag{5.10}$$

State $\Omega_l(l = Kl6)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p=1)$ is defined:

$$x_i(6,1,\mu) = -0.75x_{i-1}(6,1,\mu) + 0.61x_{i-2}(6,1,\mu) + v_i(i = 1,...,n) \tag{5.11}$$

, its spectral density function $S_{Kl6}(f_i)$ is defined

$$S_{Kl6}(f_i)=\frac{1}{11025}\frac{2}{\left|1+0.75^{\frac{-j2\pi f_i}{11025}}-0.61e^{\frac{-j4\pi f_i}{11025}}\right|^2} \tag{5.12}$$

State $\Omega_l(l=Kl3)$ sample $x_i(l,p,\mu)$ in the environment $A_p(p=1)$ is defined**:**

$$x_i(7,1,\mu)=-0.6x_{i-1}(7,1,\mu)+0.65x_{i-2}(7,1,\mu)+v_i(i=1,...,n) \tag{5.13}$$

, its spectral density function $S_{Kl7}(f_i)$ is defined

$$S_{Kl7}(f_i)=\frac{1}{11025}\frac{2}{\left|1+0.6^{\frac{-j2\pi f_i}{11025}}-0.65e^{\frac{-j4\pi f_i}{11025}}\right|^2} \tag{5.14}$$

States' functioning in the environment $A_p(p=1)$ samples $x_i(l,p,\mu)$ spectral density functions $S_l(f_i)$, when $l=Kl1,Kl2,Kl3,Kl4,Kl5,Kl6,Kl7$ are given in Figure 5.1.

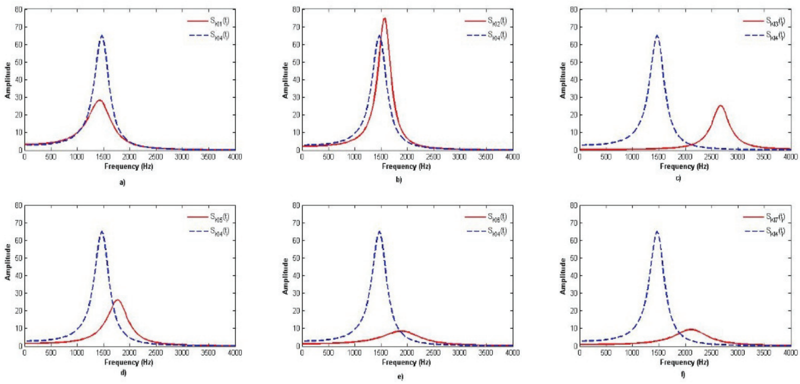

Figure 5.1 States' $\Omega_l(l=Kl1,Kl2,Kl3,Kl4,Kl5,Kl6,Kl7)$ functions in the environment $A_p(p=1)$ spectral density function comparison. The following sample $x_i(l,p,\mu)$ pairs are compared: a) $\Omega_l(l=Kl4,Kl1)$, b) $\Omega_l(l=Kl4,Kl2)$, c) $\Omega_l(l=Kl4,Kl3)$, d) $\Omega_l(l=Kl4,Kl5)$, e) $\Omega_l(l=Kl4,Kl6)$, f) $\Omega_l(l=Kl4,Kl7)$

State $\Omega_l(l = Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ samples $x_i(l, p, \mu)$ functioning in the environment $A_p(p = 2)$ are defined in equations:

State $\Omega_l(l = Kl1)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p = 2)$ is defined

$$x_i(1,2,\mu) = \frac{-0.995v_i + 0.82v_{i-1}}{-1.99x_{i-1}(1,2,\mu) + 2.592x_{i-2}(1,2,\mu) - 1.663x_{i-3}(1,2,\mu) + 0.666x_{i-4}(1,2,\mu)} \tag{5.15}$$

, and its spectral density function $S_{Kl1}(f_i)$ is defined

$$S_{Kl1}(f_i) = \frac{1}{11025} \frac{4}{\left|1 + 1.99^{\frac{-j2\pi f_i}{11025}} - 2.592e^{\frac{-j4\pi f_i}{11025}} + 1.663e^{\frac{-j6\pi f_i}{11025}} - 0{,}666e^{\frac{-j8\pi f_i}{11025}}\right|^2} \tag{5.16}$$

State $\Omega_l(l = Kl1)$ sample $x_i(l, p, \mu)$ modulations in the environment $A_p(p = 2)$ are given in Figure 5.2.

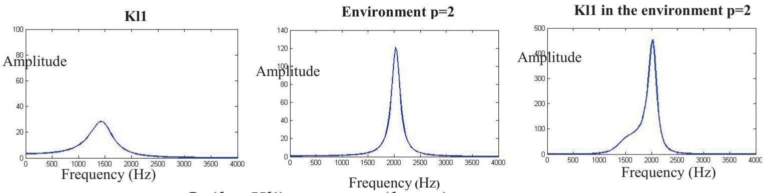

Figure 5.2 State $\Omega_l(l = Kl1)$ sample $x_i(l, p, \mu)$ modulations in the environment $p = 2$

State $\Omega_l(l = Kl2)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p = 2)$ is defined:

$$x_i(2,2,\mu) = \frac{-0{,}99v_i + 0.875v_{i-1}}{-1.98x_{i-1}(2,2,\mu) + 2.694x_{i-2}(2,2,\mu) - 1{,}742x_{i-3}(2,2,\mu) + 0.765x_{i-4}(2,2,\mu)} \tag{5.17}$$

, and its spectral density function $S_{Kl2}(f_i)$ is defined.

$$S_{Kl2}(f_i) = \frac{1}{11025} \frac{4}{\left|1 + 1.98^{\frac{-j2\pi f_i}{11025}} - 2.694e^{\frac{-j4\pi f_i}{11025}} + 1{,}7423e^{\frac{-j6\pi f_i}{11025}} - 0{,}765e^{\frac{-j8\pi f_i}{11025}}\right|^2} \tag{5.18}$$

State $\Omega_l(l = Kl4)$ sample $x_i(l, p, \mu)$ modulations in the environment $p = 2$ are given in Figure 5.3.

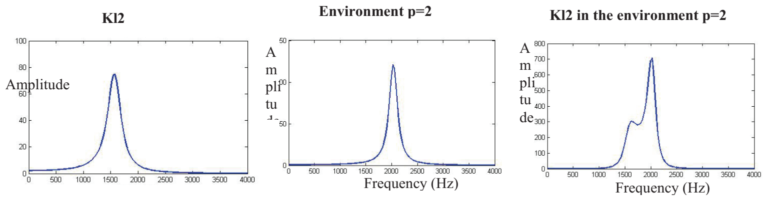

Figure 5.3 State $\Omega_l(l = Kl2)$ sample $x_i(l, p, \mu)$ modulations in the environment $p = 2$

State $\Omega_l(l = Kl3)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p = 2)$ is defined:

$$x_i(3,2,\mu) = \frac{-0.4715v_i + 0.8505v_{i-1}}{-0.943x_{i-1}(3,2,\mu) + 1.8154x_{i-2}(3,2,\mu) - 0.7695x_{i-3}(3,2,\mu) + 0.7209x_{i-4}(3,2,\mu)} \tag{5.19}$$

, and its spectral density function $S_{Kl3}(f_i)$ is defined:

$$S_{Kl3}(f_i) = \frac{1}{11025} \frac{4}{\left|1 + 0.943^{\frac{-j2\pi f_i}{11025}} - 1.8145e^{\frac{-j4\pi f_i}{11025}} + 0.7695e^{\frac{-j6\pi f_i}{11025}} - 0.7209e^{\frac{-j8\pi f_i}{11025}}\right|^2} \tag{5.20}$$

State $\Omega_l(l = Kl3)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p = 2)$ are given in Figure 5.4.

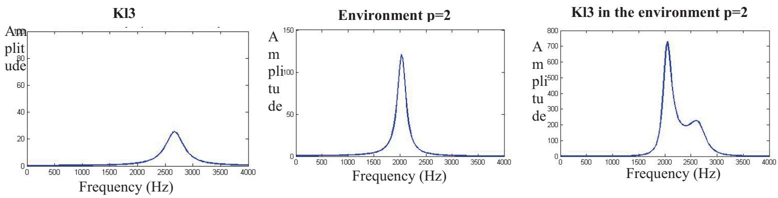

Figure 5.4 State $\Omega_l(l = Kl3)$ sample $x_i(l, p, \mu)$ modulations in the environment $A_p(p = 2)$

State $\Omega_l(l = Kl4)$ sample $x_i(l, p, \mu)$ in the environment $A_p(p = 2)$ is defined:

$$x_i(4,2,\mu) = \frac{-1.02v_i + 0.8655v_{i-1}}{-2.04x_{i-1}(4,2,\mu) + 2.723(4,2,\mu)x_{i-2}(4,2,\mu) + 1.7808x_{i-3}(4,2,\mu) + 0.7479x_{i-4}(4,2,\mu)} \tag{5.21}$$

, and its spectral density function $S_{Kl4}(f_i)$ is defined.

$$S_{Kl4}(f_i)=\frac{1}{11025}\frac{4}{\left|1+2.04^{\frac{-j2\pi f_i}{11025}}-2.723e^{\frac{-j4\pi f_i}{11025}}+1.7808e^{\frac{-j6\pi f_i}{11025}}-0.7479e^{\frac{-j8\pi f_i}{11025}}\right|^2} \qquad (5.22)$$

State $\Omega_l(l=Kl4)$ sample $x_i(l,p,\mu)$ modulations in the environment $A_p(p=2)$ are given in Figure 5.5.

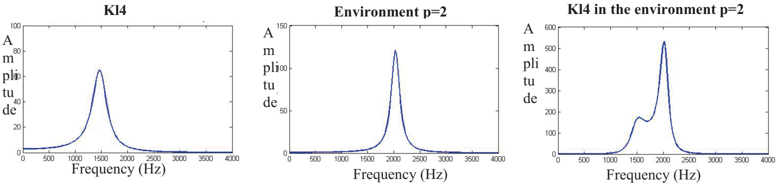

Figure 5.5 State $\Omega_l(l=Kl4)$ sample $x_i(l,p,\mu)$ modulations in the environment $p=2$

State $\Omega_l(l=Kl5)$ sample $x_i(l,p,\mu)$ in the environment $A_p(p=2)$ is defined:

$$x_i(5,2,\mu)=\frac{-0{,}8785v_i+0{,}8335v_{i-1}}{-1.7557x_{i-1}(5,2,\mu)+2.4326x_{i-2}(5,2,\mu)-1.4749(5,2,\mu)x_{i-3}(5,2,\mu)+0.6903x_{i-4}(5,2,\mu)} \qquad (5.23)$$

, and its spectral density function $S_{Kl5}(f_i)$ is defined.

$$S_{Kl5}(f_i)=\frac{1}{11025}\frac{4}{\left|1+1.7557^{\frac{-j2\pi f_i}{11025}}-2.4326e^{\frac{-j4\pi f_i}{11025}}+1.4749e^{\frac{-j6\pi f_i}{11025}}-0.6903e^{\frac{-j8\pi f_i}{11025}}\right|^2} \qquad (5.24)$$

State $\Omega_l(l=Kl5)$ sample $x_i(l,p,\mu)$ modulations in the environment $A_p(p=2)$ are given in Figure 5.6.

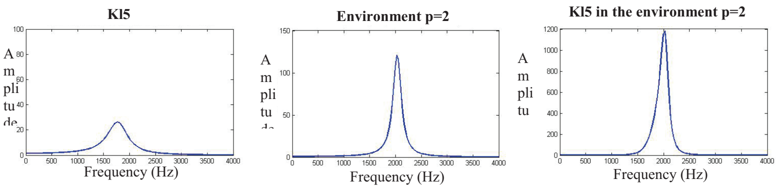

Figure 5.6 State $\Omega_l(l=Kl5)$ sample $x_i(l,p,\mu)$ modulations in the environment $A_p(p=2)$

State $\Omega_l(l=Kl6)$ sample $x_i(l,p,\mu)$ in the environment $A_p(p=2)$ is defined:

$$x(6,2,\mu) = \frac{-0,775 v_i + 0,755 v_{i-1}}{-1.55 x_{i-1}(6,2,\mu) + 2.11 x_{i-2}(6,2,\mu) - 1.163 x_{i-3}(6,2,\mu) + 0.549 x_{i-4}(6,2,\mu)} \tag{5.25}$$

, and its spectral density function $S_{Kl6}(f_i)$ is defined.

$$S_{Kl6}(f_i) = \frac{1}{11025} \frac{4}{\left| 1 + 1.55^{\frac{-j2\pi f_i}{11025}} - 2.11 e^{\frac{-j4\pi f_i}{11025}} + 1.163 e^{\frac{-j6\pi f_i}{11025}} - 0.549 e^{\frac{-j8\pi f_i}{11025}} \right|^2} \tag{5.26}$$

State $\Omega_l(l = Kl6)$ sample $x_i(l,p,\mu)$ modulations in the environment $A_p(p=2)$ are given in Figure 5.7.

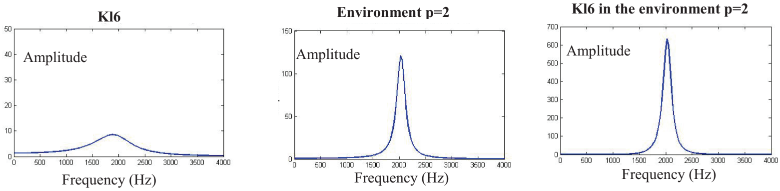

Figure 5.7 State $\Omega_l(l = Kl6)$ sample $x_i(l,p,\mu)$ modulations in the environment $A_p(p=2)$

State $\Omega_l(l = Kl7)$ sample $x_i(l,p,\mu)$ in the environmetn $A_p(p=2)$ is defined:

$$x_i(7,2,\mu) = \frac{-0,7 v_i + 0,775 v_{i-1}}{-1.4 x_{i-1}(7,2,\mu) + 2.03 x_{i-2}(7,2,\mu) - 1.106 x_{i-3}(7,2,\mu) + 0.598 x_{i-4}(7,2,\mu)} \tag{5.27}$$

, and its spectral density function $S_{Kl7}(f_i)$ is defined.

$$S_{Kl7}(f_i) = \frac{1}{11025} \frac{4}{\left| 1 + 1.4^{\frac{-j2\pi f_i}{11025}} - 2.03 e^{\frac{-j4\pi f_i}{11025}} + 1.106 e^{\frac{-j6\pi f_i}{11025}} - 0.598 e^{\frac{-j8\pi f_i}{11025}} \right|^2} \tag{5.28}$$

State $\Omega_l(l = Kl7)$ sample $x_i(l,p,\mu)$ modulations in the environment $A_p(p=2)$ are given in Figure 5.8.

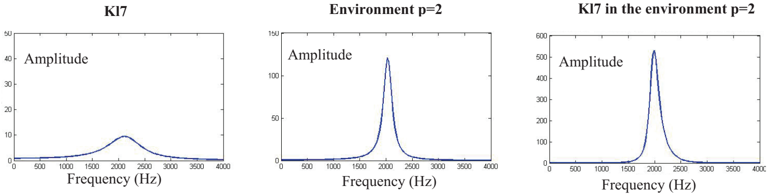

Figure 5.8 State $\Omega_l(l = Kl7)$ sample $x_i(l, p, \mu)$ modulations in the environment $A_p(p = 2)$

States $\Omega_l(l = Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ functioning in the environment $A_p(p = 2)$ samples' $x_i(l, p, \mu)$ spectral density functions $S_l(f_i)$, $l = Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7$ are given in Figure 5.9.

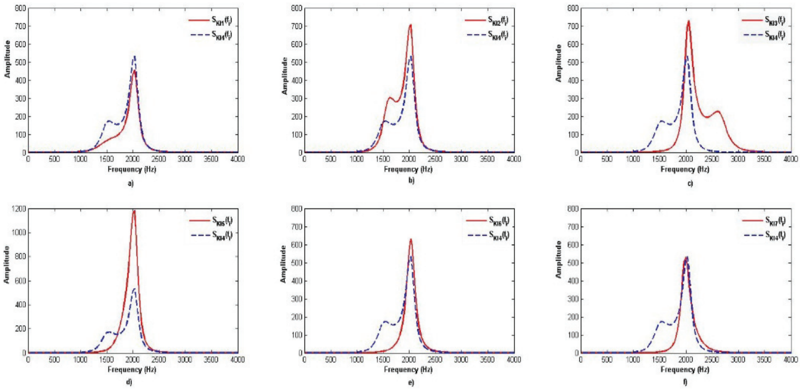

Figure 5.9 pav States $\Omega_l(l = Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ functioning in the environment $A_p(p = 2)$ spectral density functions' comparison. The following state samples' $x_i(l, p, \mu)$ pairs: a) $\Omega_l(l = Kl4, Kl1)$, b) $\Omega_l(l = Kl4, Kl2)$, c) $\Omega_l(l = Kl4, Kl3)$, d) $\Omega_l(l = Kl4, Kl5)$, e) $\Omega_l(l = Kl4, Kl6)$, f) $\Omega_l(l = Kl4, Kl7)$ are compared

5.2 *Research on methods evaluating the moment of change in the environment conditions*

While analyzing problems of acoustic speech signals', or signals' of any other kind recognition device adaptation to changing conditions, it is important to determine the moment u correctly and in due course - it states when the situation observed undergoes a change basing on which the recognition device can start conducting adaptation to newly-created conditions.

The aim

The aim of the work is to elucidate which of the methods described in section 4.2 would be most appropriate to use for solving the problem analyzed willing to correctly and in due course establish time moment *u* when the recognition device shifts from its primar environment $A_p(p=1)$,in which it was trained to recognize state $\Omega_l(l=Kl1, Kl2, Kl3, Kl5, Kl6, Kl7)$ generated samples $x_i(l,p,\mu)$, to new environment $A_p(p=2,3,..,P)$. The device needs to start operating the recognition process willing to continue recognition procedure in the new environment $A_p(p=2,3,..,P)$ retaining the quality $P(x_i(l,p,\mu))$ of recognition (Figure 4.11) demonstrated in the primal environment $A_p(p=1)$.

Proceeding of the research

State $\Omega_l(l=Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ samples $x_i(l,p,\mu)$ described in section 5.1 are going to be used in the research. State $\Omega_l(l=Kl4)$ samples $x_i(l,p,\mu)$ will be compared to state $\Omega_l(l=Kl1, Kl2, Kl3, Kl5, Kl6, Kl7)$ samples $x_i(l,p,\mu)$ - state $\Omega_l(l=Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ samples' spectral density function being $S_l(f_i)$, when $l=Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7$ given in Figure 5.3 – in it *l* describes the state of sample, *p* describes environment in which the state generates the aforementioned samples, μ-decribes the rank number $\mu=1,2....$ of the generated sample.

A situation in which, during the primal moment of recognition, the recognition device is trained to recognize state $\Omega_l(l=Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ samples $x_i(l,p,\mu)$ functioning in the environment $A_p(p=1)$ and at a particular time moment *u*, the recognition device enters environment $A_p(p=2)$, is modeled. The research is to determine, which of condition change moment evaluating methods discussed in section 4.2 would be most efficient in solving the analyzed problem of adapting to changing conditions. According to the specifics of operation, the following methods recording the moment of signal change have been selected:

- Shewhart method
- Moving average
- Momentary specific occasional elements' recognition (TSEA)

Figure 5.10 illustrates parameter α_i estimate sequence acquired during state $\Omega_l(l=Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ samples $x_i(l,p,\mu)$ recognition (the parameter is discussed in more detain in chapter 3). In the sequence, the methods analyzed are searching for the environment conditions change moment *u*. Values of the parameter α_i during the dynamic system's functioning in environment $A_p(p=1)$ are given in the yellow background. Values of the parameter α_i during the dynamic system's functioning in environment $A_p(p=2)$ are given in the white background.

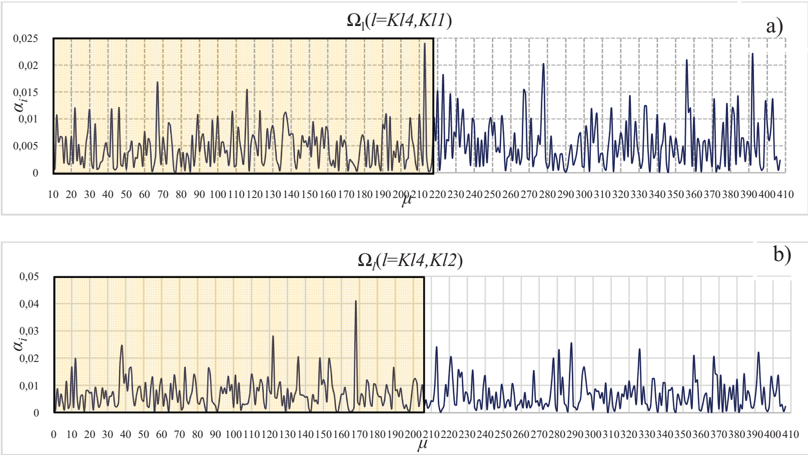

Figure 5.10.1. Values of the parameter α_i acquired during recognition process of states $\Omega_l(l = Kl1, Kl2, Kl4)$ functioning in environment $A_p(p = 1)$ are given in the yellow background; the white background illustrates values acquired in environment $A_p(p = 2)$. Compared are the following state sample $x_i(l, p, \mu)$ pairs: a) $\Omega_l(l = Kl4, Kl1)$, b) $\Omega_l(l = Kl4, Kl2)$

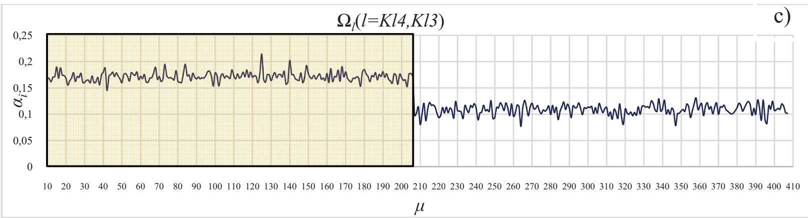

Figure 5.10.2. Values of the parameter α_i acquired during recognition process of states $\Omega_l(l = Kl3, Kl4)$ functioning in environment $A_p(p = 1)$ are given in the yellow background; the white background illustrates values acquired in environment $A_p(p = 2)$.

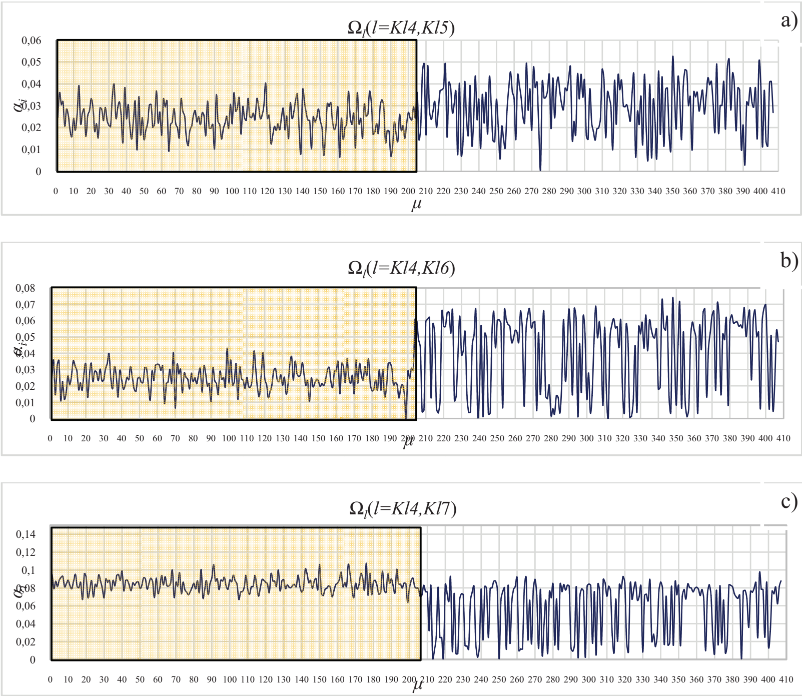

Figure 5.10.3. Values of the parameter α_i acquired during recognition process of states $\Omega_l(l = Kl4, Kl5, Kl6, Kl7)$ functioning in environment $A_p(p=1)$ are given in the yellow background; the white background illustrates values acquired in environment $A_p(p=2)$. Compared are the following state sample $x_i(l,p,\mu)$ pairs: a) $\Omega_l(l = Kl4, Kl5)$,b) $\Omega_l(l = Kl4, Kl6)$, c) $\Omega_l(l = Kl4, Kl7)$

Results

Figure 5.11 illustrates the results of the research that demonstrate which of the methods analyzed consumed the least amount of time while establishing the fact of change in the primal conditions of state $\Omega_l(l = Kl1, Kl2, Kl3, Kl4, Kl5, Kl6, Kl7)$ samples $x_i(l,p,\mu)$ observed. This parameter is highly important as when the analysis takes place in the real situation and real man-uttered sounds are examined, the recognition device cannot claim for numerous repetitions of particular sounds, thus, the speed of identifying new environment is also a very important characteristic.

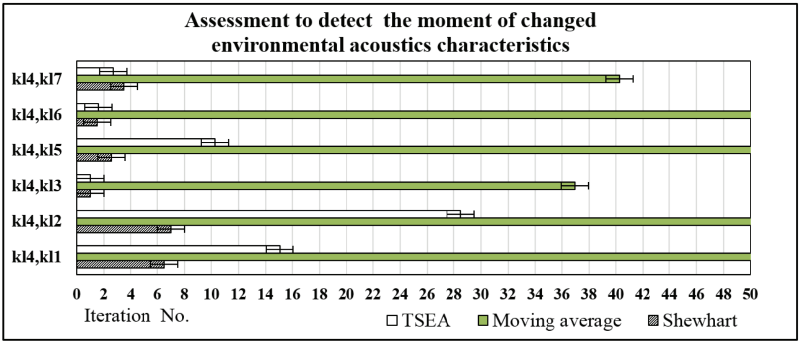

Figure 5.11 Analysis of detecting a moment of change in enviromental conditions

As it is evident from the data given in **Figure 5.11, *Schewhart*** method was least time-consuming when identifying the moment of change in the conditions of the environment. ***The moving average*** method demonstrated the weakest results as using it results in either a very slow establishment of the conditions moment, or it does not cope with the task being unable to determine the moment. Such results have been obtained in cases $\Omega_l(l = Kl4, Kl6)$, $\Omega_l(l = Kl4, Kl5)$, $\Omega_l(l = Kl4, Kl1)$ and $\Omega_l(l = Kl4, Kl2)$. Average results have been demonstrated by the method of momentary specific occasional occasional elements' recognition – it is wise to use this method when state samples carry different characteristics, or their characteristics are of average similarity.

Conclusions

Three methods evaluating the change moment of environment conditions have been analyzed during the experimental research. A detailed overview of the methods' operation is given in chapter 4.2. Experimental results acquired have shown that ***Scewhart*** method is the most suitable for solving the analyzed problem of adapting to changed conditions. This method is fastest in establishing a fact that a change has occurred in the environment observed. Moreover, methodology of this method's operation is easily adaptable to newly-formed recognition situation.

5.3 Research on the effectiveness of methods for adapting to changed environment conditions

While analyzing the recognition system's adaptation to changed recognition conditions, it is of utmost importance to assess all possible adaptive factors. It is necessary to know under what conditions adapting to new circumstances is very difficult, and under what conditions adaptation is always possible.

In order to retain the accuracy of the research as high as possible, artificial model signals are going to be used. These allow the modeling of state elements possessing diverse characteristics.

The aim

The aim of the experimental research is to analyze the effectiveness of the methods APMIA, NMPT, ELLIPSE observing state samples possessing different features.

Proceedings of the research

The following tasks have to be performed when conducting the experimental research:

- Elements of the observed state possessing various characteristics are modeled.
- The recognition device is trained to recognize the elements of the observed states.
- Features characteristic to environments in which the research is going to be carried out are defined. Primal parameters are given to the recognition device.

Adaptation situations

The aim of the problem analyzed is to find a method which would help the recognition device to adapt to changing acoustic conditions of the environment. Four adaptation situations are going to be analyzed. In all four situations the recognition device is at first trained to recognize samples generated by two states in environment $A_p(p=1)$ possessing certain acoustic qualities. Then, at a certain time moment u, the recognition device moves from environment $A_p(p=1)$ to environment $A_p(p=2)$ which holds different acoustic characteristics.

Results

Results acquired during the experimental research are given in Figure 5.12, Figure 5.13 and Figure 5.14.

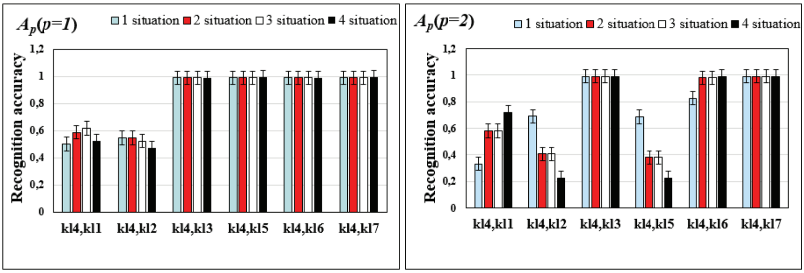

Figure 5.12 Recognition results using APMIA method

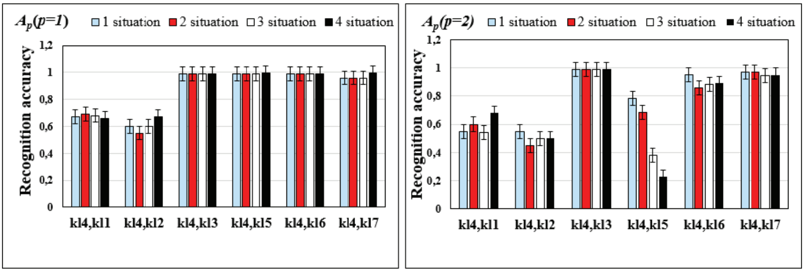

Figure 5.13 Recognition results using NMPT method

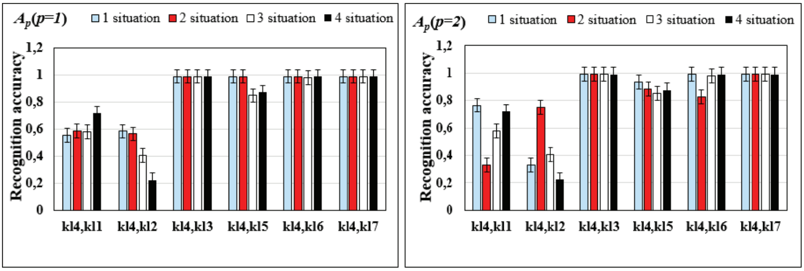

Figure 5.14 Recognition results using ELIPSE method

The acquired results and situations' evaluation analysis have shown that it is most effective to use different adaptation methods when the level of state similarity is different.

- When the distance between classes is small:
  - 3rd method of adaptation (Ellipse) and 1st method of adaptation (APMIA) are applicable to situations 1 and 2.
  - 3rd method of adaptation (Ellipse) is applicable to situations 3 and 4.
- When the distance between classes is of average length:
  - 3rd method of adaptation (Ellipse) is applicable to situation 1.
  - 1st method of adaptation (APMIA) is applicable to situation 2.
  - 1st method of adaptation (APMIA) and 3rd method of adaptation (Ellipse) are applicable to situations 3 and 4.
- When the distance between the classes is large:
  - All methods of adaptation are acceptable.

Conclusions

Results of the experimental research conducted have shown that recognition quality improves when the adaptation methods suggested are used. Results of the research have defined the factors influencing the quality of recognition: the sequence of state samples' emergence in the environment, order and similarity of the state samples' features.

Recognition quality cannot be improved if features of the state sample observed are very similar, i.e. if the distance difference acquired during recognition is equal, or nearly equal, to 0; also, if features of state samples become very similar after entering a new environment.

If during recognition the state samples observed emerge rarely, it forecasts a very difficult adaptation to the new environment as the emerging state samples are too few for the training procedure to be carried out successfully. Thus, the quality of recognition might worsen.

RESULTS AND CONCLUSIONS

In the dissertation, the following results have been obtained while solving a problem of adapting acoustic signals' recognition systems to altered acoustic environments:

1. A theoretical model for adapting acoustic signal recognizer to changed closed acoustic spaces has been created together with accompanying constructive methods.
2. Methods for evaluating acoustic characteristics of environments in which the training process of acoustic signals' recognizer, and acoustic signals' recognition take place have been investigated.
3. Factors of closed acoustic environments influencing the functioning of acoustic signal recognizers have been analyzed.
4. A method using the acoustic signals' recognition results for establishing the moment of change in the acoustic environments has been proposed.
5. Methods evaluating the moment of change in the environment conditions have been analyzed.
6. Procedures of recognizers' control and functioning allowing the devices' adaptation to altered conditions of operation have been created and investigated.
7. Experimental database with the help of which investigation of various acoustic signals' recognition situations is possible through adaptation to changed conditions of acoustic environments, has been created.
8. Practical adaptability of the experimental database created has been demonstrated.

Information about the author

Gintarė Čeidaitė was born in Jurbarkas, 2 April, 1982.

She was granted a Bachelor's Degree in Informatics at the Faculty of Informatics of Vytautas Magnus University in 2005 and a Master's Degree in Applied Informatics at the Faculty of Informatics of Vytautas Magnus University in 2007. From 2007 she works as assistant in Department of Systems' Analysis, Faculty of informatics, Vytautas Magnus University. In 2007–2014 she was a PhD student of Vytautas Magnus University.

E-mail: g.ceidaite@if.vdu.lt

References

[1] **Caballero A. F., Sanfeliu A., Shira Y**. Pattern Recognition in Interdisciplinary Perception and Intelligence (Internet address: : http://upcommons.upc.edu/e-prints/bitstream/2117/2700/1/doc1.pdf, žr. 2014 02 15).

[2] **Bishop M. C**. Pattern Recognition and Machine Learning 2006 Springer 738psl

[3] **Deller J. R, John H. L. H., Proakis G**. Discrete-Time Processing of Speech Signals // The Institute of Electrical and Electronics Engineers, lnc., NewYork, 1993..

[4] **Wolfel M., McDonough J.** Distant Speech Recognition, 2009, John Wiley & Sons Ltd, 573 p.

[5] **Ruzgys G., Kamarauskas J**. Nuo teksto nepriklausomas kalbančiojo atpažinimas. 14-osios Lietuvos jaunųjų mokslininkų konferencijos „Mokslas – Lietuvos ateitis“, ISSN 2029-7149, online 2011 metų teminės konferencijos straipsnių rinkinys

[6] **Laurinčiukaitė S., Lipeika A.** Syllable-Phoneme based Continuous Speech Recognition // Electronics and Electrical Engineering 2006. Nr. 6(70) 91–94 p.

[7] **Laurinčiukaitė S**., Lietuvių šnekos atpažinimo akustinis modeliavimas. doctoral thesis 2008.

[8] **Kamarauskas J.** Asmens atpažinimas pagal balsą. . doctoral thesis, 2008.

[9] **Raškinis G., Raškinienė D.** Building medium-vocabulary isolated-word Lithuanian HMM speech recognition system. Informatica, 14(1), 75–84. 2003.

[10] **Vaičiūnas A.** Lietuvių Kalbos statistinių modelių ir jų taikymo šnekos atpažinimui tyrimas, kai naudojami labai dideli žodynai. . doctoral thesis, 2006

[11] **Šilingas D., Raškinis G., Telksnys L**. 2004a. Review of Lithuanian speech and language processing, in Proc. of Int. Conf on Human Language Technologies – The Baltic Perspective, 144–150.

[12] **Šilingas D., Telksnys L.** Specifics of hidden Markov model modifications for large vocabulary continuous speech recognition // Informatica 15(1), 2004, 93–110 p.

[13] **Šilingas D.** Choosing acoustic modeling units for Lithuanian continuous speech recognition based on hidden Markov models, . doctoral thesis, Kaunas, 2005. 162 p

[14] **Vaičiūnas A., Raškinis G.** Statistinių lietuvių kalbos modelių kūrimas ir pirminis tyrimas. Konferencijos „Informacinės technologijos 2003“ pranešimų medžiaga, IX–(35–40).

[15] **Vaičiūnas A., Raškinis G., Kaminskas V.**. Statistical language models of Lithuanian based on word clustering and morphological decomposition // Informatica 15(4): 2004. 565–580 p.

[16] **Šalna B., Kamarauskas J.** Evaluation of Effectiveness of Different Methods in Speaker Recognition // Electronics and Electrical Engineering, 2010, Nr. 2(98), 67–70 p.

[17] **Rudžionis A., Ratkevičius K., Rudžionis V.** Speech in Call and Web centers // Electronics and Electrical Engineering. Kaunas: Technologija, 2005. No. 3(59). 58–63 p.

[18] **Maskeliūnas R.** Lietuviškų balso komandų atpažinimas daugybinių transkripcijų pagrindu. Daktaro disertacija, 2009.

[19] **Maskeliūnas R., Rudžionis A., Ratkevičius, K., Rudžionis V.** Investigation of Foreign Languages Models for Lithuanian Speech Recognition // Electronics and Electrical Engineering, 2009, vol. 3 (91), 15–21 p.

[20] **Rudžionis A., Ratkevičius K., Rudžionis V., Kasparaitis P**. An Experience of Creating Lithuanian Speech-Enabled Web Applications. In: *The second Baltic conference on Human Language Technologies*, Proceedings, Tallinn, Estonia, April 4–5, 2005, 335-340. ISBN 9985-894-83-9.

[21] **Rudžionis V., Ratkevičius K., Rudžionis A., Raškinis G.,** Recognition of Voice Commands Using Hybrid Approach. Springer-Verlag CCIS (Communications in Computer and Information Science), Volume 403, 2013, 249–260 p. ISBN 978-3-642-41946-1.

[22] **Lipeika A., Lipeikienė J.** Word Endpoint Detection Using Dynamic Programming // Informatica, 2003, 14(4), 487–496 p.

[23] **Tamuliavičius G.** Pavienių žodžių atpažinimo sistemų kūrimas. . Doctoral thesis, 2008

[24] **Lipeika A., Lipeikienė J., Telksnys L.** Development of Isolated Word Speech Recognition System // Informatica 2002, Vol. 13, No. 1, 37–46 p.

[25] **Lileikytė R.** Šnekos atpažinimo požymių kokybės vertinimas. Doctoral thesis, 2012.

[26] **Lileikytė R., Telksnys L**. Quality Estimation Methodology of Speech Recognition Features // Electronics and Electrical Engineering, 4(110), 113–116 p. 2011. ISSN 1392–1215.2

[27] **Telksnys. L**. Atsitiktinių procesų savybių lėto kitimo aptikimas. Statistinės valdymo problemos. Atsitiktinių procesų atpažinimo metodai. 89 laida, Vilnius, 1990. Lietuvos MA Matematikos ir kibernetikos institutas. 86–93 p.

[28] **Stahl V., Res P**. Lab., Aachen, Germany; **Fischer, A., Bippus, R.** Acoustic synthesis of training data for speech recognition in living room environments // Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001, IEEE International Conference (Volume: 1), 21–24 p.

[29] **Okawa S., Bocchieri E., Potamianos** A. Multi-band speech recognition in noisy environments // Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference (Volume: 2), 641–644 p.

[30] **Moreno P. J.** Speech Recognition in Noisy Environments. . Doctoral thesis, 1996.

[31] **Das S., Bakis R., NBdas A., Nahamoo D., Picheny M.** Influence of background noise and microphone on the performance of the IBM TANGORA speech recognition system, 1993, IEEE.

[32] **Denda Y., Nishiura T., Kawahara H**. Noisy speech recognition with microphone array steering and fourier/wavelet spectral subtraction. Faculty of Systems Engineering, Wakayama University, 28 Sept.-1 Oct. 2003, IEEE.

[33] **Francois H., Pearce D., Rex J.** Dual-microphone robust front end for arm‘s length speech recognition in 2006 International Workshop on Acoustic Echo and Noise Control.

[34] **Wan E. A., Nelson W. A. T.** Removal of noise from speech using the dual EKF algorithmt // Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on (Volume: 1), 381–384 p.

[35] **Verteletskaya E., Simak B.** Speech Distortion Minimized Noise Reduction Algorithm // Proceedings of the World Congress on Engineering and Computer Science, 2010, Vol I WCECS 2010, October 20–22, 2010, San Francisco, USA.

[36] **Jiang W., Malvar H.** Adaptive noise reduction of speech signals, July 2000.

[37] **Basseville M., Nikiforov I. V.** Detection of Abrupt Changes: Theory and Application ISBN 0-13-126780-9, April 1993.

[38] **Khoo M. B. C.** A Moving Average Control Chart for Monitoring the Fraction Non-conforming // Quality and reliability engineering international, 2004, 20: 617–635 p.

[39] Handbook for basic process improvement. Control chart (Internet address:

http://www.au.af.mil/au/awc/awcgate/navy/bpi_manual/mod10-control.pdf, accessed. 2013 12 15).

[40] **Frank P.** Control charts in quality control –shewart charts application

(Internet address: http://www.feec.vutbr.cz/EEICT/2003/fsbornik/03-PGS/04-Power_Electrical_Engineering/09-frank_petr.pdf, accessed. 2013 12 15).

[41] **Kruopis J.** Matematinė statistika, 1977, Vilnius..

[42] **Telksnys L., Kaukėnas J.** Recognition of Short-Time Specific Random Elements in Random Sequences.

[43] **Karčiauskas K.** Antros eiles kreivių ir paviršių geometrija (Internet address::

http://www.techmat.vgtu.lt/konspektai/Analizine_geom/Antros%20Eiles%20Kreiviu%20ir%20Pavirsiu%20Geometrija%20(K.%20Karciauskas).pdf, accessed.. 2013 11 23).

[44] **Stauskis V., Kunigėlis V.** Įvairių tipų impulsinių garso šaltinių akustikos charakteristikos. Statyba, Civil engineering, 1998. IV tomas Nr. 4.

Gintarė ČEIDAITĖ

AKUSTINIŲ SIGNALŲ ATPAŽINIMO SISTEMŲ PRISITAIKYMO PRIE PASIKEITUSIŲ APLINKOS SĄLYGŲ TYRIMAS

Santrauka

Tyrimų sritis ir problemos aktualumas

Šiomis dienomis IT technologijos sparčiai tobulėja, pasiūlydamos žmonėms vis naujesnių funkcionalumų, galimybių. Siekiama, kad IT technologijos būtų kuo draugiškesnės, patrauklesnės paprastam žmogui. Taip sparčiai progresuojant, galima prognozuoti, kad jos greitai bus ne tik mobilios, valdomos vos kelių pirštų pagalba, bet turės galimybę būti valdomos balsu.

Technologijų valdymas balsu, tolimesnio progreso perspektyva. Su tokiomis technologijomis, ar tai būtų nešiojami, planšetiniai kompiuteriai, išmanūs mobilieji telefonai, ar koks nors kitas mobilus įrenginys. Žmonės juos naudos visur, visiškai nesvarbu kur jie bebūtų. Ar tai būtų kavinė, autobuso stotis, metro, ar lietus lytų, ar tai būtų panaši labai triukšminga aplinka. Technologijos privalės sėkmingai funkcionuoti esant įvairioms sąlygoms. Tad labai svarbu išnagrinėti įvairių triukšmų daroma poveikį įrenginyje esančiam atpažinimo įrenginiui ir galimybes , juos eliminuoti, bei adaptuotis prie esamos situacijos.

Disertacijoje nagrinėjama garso signalų atpažintuvų funkcionavimo nestacionariose akustinėse aplinkose problema. Pateikiama teorija ir konstruktyvūs metodai ***vertinimui uždarų patalpų akustinių savybių***, sukeliančių garso signalų atpažinimo klaidas bei ***garso signalų atpažinimo tikslumo didinimui***, atpažintuvams patekus į patalpas su skirtingomis akustinėmis savybėmis.

Analizuojama aplinkos akustinių savybių daroma įtaką atpažinimo įrenginio atpažinimo kokybei. Disertacijoje pateikiamas adaptavimosi prie pakitusių sąlygų algoritmas. Kurio pagalba galima laiku įvertinti aplinkos savybių pasikeitimo momentą, ir mokymo procedūrą adaptuoti, prie naujų sąlygų.

Kadangi nagrinėjami stochastiniai signalai, tad eliminavus aplinkos poveikį, arba jį prislopinus, galima prarasti svarbios stebimame signale esančios informacijos. Tad vienintelis geriausias sprendimas šioje situacijoje, būtų adaptuotis prie naujų sąlygų. Permokyti atpažinimo įrenginį, kad jis sugebėtų atpažinti stebimas būsenos realizacijas apmokytas ne tik konkrečioje aplinkoje, bet ir sugebėtų apsimokyti esant naujoje aplinkoje.

Disertacijoje pateikiamas teorinis situacijos modelis, kuriame nurodoma, kaip būtų galima aprašyti kiekvieną spendžiamo uždavinio sudedamąją dalį, t.y. kaip modeliuojami akustinėmis savybėmis pasižymintys signalai, kaip įvertinamos ir modeliuojamos aplinkos akustinės savybės, kurioje atpažinimo įrenginys funkcionuoja, kaip pasikeičia šnekos signalo savybės papuolus, į naujomis akustinėmis savybėmis pasižyminčia aplinką. Bei aprašomi , kokie papildomi veiksniai įtakoja, ar gali įtakoti atpažinimo kokybę.

Eksperimentiniams tyrimams atlikti, buvo sukurta dinaminio laiko kraipymu metodu paremta programinė įranga, kurios pagalba, galima modeliuoti įvairias atpažinimo situacijas. Valdant ne tik stebimų būsenų realizacijas, bet ir aplinkos, kurioje būsena funkcionuoja savybes.

Atlikti eksperimentinius tyrimus su realiais šnekos signalais yra labai brangu, reikalauja labai didelių žmogiškųjų ištekli. Tad disertacijoje šie akustiniai signalai dirbtinai generuojami išlaikant jų savybes. Modeliuojant įvairias galimas atpažinimo situacijas, bei būsenų generuojamas realizacijas galima įvertinti įvairius mokymo procedūros adaptavimo momentus. Galima pasakyti, kada siūlomas modelis veiks gerai, kada adaptuotis yra sudėtinga ir kokioms sąlygoms esant siūlomas adaptavimosi modelis neveiks. Disertacijoje pateikiami eksperimentiniai tyrimų rezultatai atsakantysis šiuos klausimus .

Adaptavimosi prie kintančių aplinkos sąlygų metodo veiksmingumas, demonstruojamas disertacijoje aprašomi eksperimentiniai tyrimai.

Aprašomi adaptavimosi prie pakitusių sąlygų algoritmai remiasi atpažinimo įrenginio gautais atpažinimo rezultatais, tad šis modelis, nėra taikomas tik konkrečiam atpažinimo įrenginiui, bet gali būti naudojamas, kitus atpažinimo metodus naudojančiose atpažinimo sistemose.

Darbo tikslas ir uždaviniai

Darbo tikslas–išanalizuoti akustinių signalų atpažinimo įrenginio galimybes prisitaikyti prie kintančių aplinkos sąlygų, bei pasiūlyti konstruktyvią prisitaikymo prie kintančių aplinkos sąlygų metodiką.

Uždaviniai:

1. Aprašyti aplinkų akustines savybes, kuriose vyksta akustinių signalų atpažinimo procesai (apmokymas ir atpažinimas)
2. Išanalizuoti aplinkos akustinių savybių daromą įtaką atpažinimo įrenginiui.
3. Išanalizuoti metodus leidžiančius aptikti aplinkos savybių pasikeitimo momentą.
4. Išanalizuoti atpažinimo įrenginio mokymo procedūros valdymo galimybes pasikeitusiose aplinkose.
5. Sukurti eksperimentinę bazę, kurios pagalba galima modeliuoti įvairias akustinių signalų atpažinimo situacijas, valdant aplinkų akustines savybes.
6. Eksperimentiniais tyrimais pademonstruoti pasiūlyto atpažinimo įrenginio prisitaikymo prie kintančių aplinkos sąlygų metodo efektyvumą.

Tyrimo metodai

Teorinei analizei, praktinei realizacijai panaudotos tikimybių teorijos, matematinės statistikos, dinaminių sistemų teorijos, akustikos, skaitmeninio signalų apdorojimo, bei atpažinimo teorijos žinios.

Aplinkos akustinių savybių analizei atlikti buvo panaudota, patalpų akustinių charakteristikų vertinimo metodika. Patalpų akustinių charakteristikų vertinimui, buvo naudojamas impulsinis akustis signalas kurį sukelia 9 mm kalibro revolveris. Jo sukelto impulsinio signalo atsakas buvo perkeliamas į kompiuterį naudojant CoolEdit programinės įrangos paketą. Gauto signalo spektrinės charakteristikos skaičiuojamos naudojant STADIA programinę įrangą.

Patalpų akustinių signalų daromai įtakai nustatyti, buvo panaudotas dinaminio laiko kraipymo metodu paremtas atpažinimo įrenginys *Atpažintuvas* [23] diktorių ištarti signalai apdorojami panaudojant CoolEdit programinį paketą. Tyrimams atlikti buvo surinkta 100 įvairių amžiaus kalbėtojų šnekos signalų pavyzdžių, 10 skirtingose akustinėse savybėmis pasižyminčiose aplinkose.

Aplinkos akustinių signalų, stebimų signalų, atpažinimo situacijų modeliavimui ir aplinkos savybių pasikeitimo momentui nustatyti, buvo sukurta Matlab aplinkoje veikianti programinė įranga.

Akustinių signalų atpažinimui buvo naudojamas dinaminio laiko skalės kraipymo metodu paremtas atpažinimo įrenginys.

Mokslinis naujumas

1. Sukurtas akustinių signalų atpažinimo sistemos prisitaikymo prie naujų pakitusių atpažinimo sąlygų metodas.
2. Padaryta programinės įrangos sistema, kurios pagalba galima tyrinėti įvairias akustinių signalų atpažinimo situacijas.
3. Atlikti eksperimentiniai tyrimai patvirtinantys sukurtojo akustinių signalų prisitaikymo prie kintančios aplinkos metodo naudingumą

Praktinė darbo reikšmė

Disertacijoje pasiūlyti metodai leidžia funkcionuojančiai atpažinimo sistemai kintant aplinkos sąlygoms prisitaikyti prie jų, išlaikant atpažinimo kokybę.

Šio darbo metodika paremta eksperimentinių tyrimų metu gautais rezultatais. Prisitaikymo prie kintančių aplinkos sąlygų metodui svarbu, atpažinimo įrenginio pateikiami rezultatai, o ne įrenginio veikimo specifika, tad metodas gali būti pritaikytas įvairius atpažinimo metodus naudojančiuose atpažinimo įrenginiuose.

Naudojant disertacijos metu sukurta programinį paketą, galima modeliuoti įvairias atpažinimo situacijas ir toliau tęsti pritaikymo prie kintančių aplinkos sąlygų tiriamuosius darbus.

Disertacijos struktūra ir turinys

Įvade aprašoma disertacijos tyrimo objektas, temos aktualumas, suformuojamas darbo tikslas, aprašyti tyrimo metodai, mokslinis naujumas, praktinė reikšmė, darbo aprobavimas, darbo publikacijų sąrašas, pristatoma darbo struktūra, turinys.

Pirmajame skyriuje apžvelgiame kokie nagrinėjama tematika darbai yra atilikti užsienio ir Lietuvos mokslininkų. Apžvelgiama šių mokslininkų atlikti darbai. Taip pat atliekamas atpažinimo įrenginį supančios aplinkos daromo įtakos atpažinimo kokybei tyrimas. Pateikiama analitinė veiksnių įtakojančių atpažinimo įrenginio atpažinimo kokybės veiksnių , bei šios klausimą nagrinėjančių darbų apžvalga. Pateikiami metodai, kurie slopina šių veiksnių įtaką, pateikiami šia tematika nagrinėjančių darbų apžvalga.

Antrajame skyriuje nagrinėjami veiksniai įtakojantys patalpų akustiką. Aprašomas metodas patalpų akustinių savybių matuoti, bei aprašyti. Pateikiamas metodas modeliuoti akustiniams signalams konkrečiomis akustinėmis savybėmis pasižyminčioje aplinkoje. Aprašomas eksperimentinis tyrimas, kurio rezultatai demonstruoja, kokią įtaką akustinių signalų atpažinimo įrenginiui turi patalpa, kurioje jis funkcionuoja.

Trečiame skyriuje analizuojamas aplinkos savybių pasikeitimo momento nustatymo klausymas. Pateikiamas parametras, kurį stebint galima nustatyti aplinkos savybių pasikeitimo momentą. Aprašoma, kokiais būdais stebimos būsenų realizacijos gali būti generuojamas aplinkoje. Pateikiami metodai, skirti aplinkos savybėms įvertinti.

Ketvirtame skyriuje pateikiamas trys konstruktyvus metodai, kurių pagalba atpažinimo įrenginys prisitaikyti ir efektyviai funkcionuoti naujoje aplinkoje. Detaliai nurodomi kiekvieno metodo veiksmai, kuriuos atlikus atpažinimo įrenginys gali prisitaikyti prie naujos susidariusios situacijos.

Penktame skyriuje aprašoma programinė įranga, kurios pagalba galima modeliuoti aplinkos, bei stebimų būsenų realizacijų savybes, modeliuoti jų pasirodymo sekas atpažinimo metu. Programinės įrangos pagalba galima modeliuoti įvairias atpažinimo situacijas ir prisitaikyti prie esamos situacijos Šios programinės įrangos pagalba buvo atliekami eksperimentiniai tyrimai demonstruojantys prisitaikymo prie kintančių sąlygų metodo efektyvumą.

Šeštame skyriuje aprašomi atlikti eksperimentiniai tyrimai. Pateikiamas detalus naudojamų būsenų realizacijų ir jų modeliavimo naujoje akustinėmis savybėmis pasižyminčioje aplinkoje aprašymas. Atliktas eksperimentinis tyrimas, kurio metu mėginama išsiaiškinti, kurį metodą aprašyta 3 skyriuje t, ikslingiausią būtų galima taikyti, siekiant kuo greičiau nustatyti aplinkos savybių pasikeitimo momentą. Pateikiami tyrimo metu gauti rezultatai ir išvados. Taip pat aprašomas eksperimentinis tyrimas analizuojantis prisitaikymo prie kintančių aplinkos sąlygų metodų efektyvumą. Analizuojami 4 skyriuje aprašomi prisitaikymo metodai, jų efektyvumas, pateikiami tyrimo rezultatai ir išvados.

Rezultatai ir išvados

1. Sukurtas akustinių signalų atpažinimo įrenginio / atpažintuvo prisitaikymo prie pasikeitusių uždarų akustinių erdvių teorinis modelis ir konstruktyvūs metodai.
2. Ištirti vertinimo metodai akustinių savybių patalpų, kuriose vyksta akustinių signalų atpažinimo įrenginio / atpažintuvo apmokymas ir akustinių signalų atpažinimas skirtingose aplinkose.
3. Išanalizuoti uždarų akustinių aplinkų veiksniai, įtakojantys akustinių signalų atpažinimo įrenginių / atpažintuvų funkcionavimą.
4. Pasiūlytas signalų atpažinimo rezultatu grindžiamas veiksnys, kuriuo remiantis galima nustatyti aplinkos savybių pasikeitimo momentą.
5. Išanalizuoti aplinkos savybių pasikeitimo momentą vertinantys metodai. Pateikiami eksperimentinio tyrimo rezultatai.
6. Ištirtos atpažinimo įrenginių / atpažintuvų valdymo funkcionavimo procedūros, atveriančios galimybes prisitaikyti dirbti prie pasikeitusių darbo sąlygų.
7. Sukurta eksperimentinė bazė, kurios pagalba galima tirti įvairias akustinių signalų atpažinimo situacijas, prisitaikant prie pakitusių akustinių aplinkų savybių.
8. Parodytas sukurtosios eksperimentinės bazės praktinis pritaikomumas.

Informacija apie autorę

Gintarė Čeidaitė gimė 1982 m. spalio mėn. 2 d. Jurbarke.
2005 m. įgijo informatikos bakalauro kvalifikaciją Vytauto Didžiojo universiteto Informatikos fakultete. 2007 m. įgijo taikomosios informatikos magistro laipsnį Vytauto Didžiojo universiteto Informatikos fakultete. Nuo 2007 dirba asistente Sistemų analizės katedroje, Informatikos fakultete, Vytauto Didžiojo universitete. 2007 – 2014 m. – VDU Informatikos fakulteto Sistemų analizės katedros doktorantė.

E-paštas: g.ceidaite@if.vdu.lt

Gintarė ČEIDAITĖ

ACCOUSTIC SIGNALS RECOGNITION EQUIPMENT ADAPTATION TO THE VARIOUS ENVIRONMENTS

Summary of Doctoral Dissertation

Išleido ir spausdino – Vytauto Didžiojo universiteto bibliotekos Leidybos skyrius
(S. Daukanto g. 27, LT-44249 Kaunas)
Užsakymo Nr. K14-059. Tiražas 40 egz. 2014 08 18.
Nemokamai.