VILNIUS UNIVERSITY

Rasa Lileikytė

QUALITY ESTIMATION OF SPEECH RECOGNITION FEATURES

Summary of Doctoral Dissertation
Technological Sciences, Informatics Engineering (07T)

Vilnius, 2012

Doctoral dissertation was prepared at Vilnius University, Institute of Mathematics and Informatics in 2007–2012.

Scientific Supervisor

Prof Dr Habil Laimutis Telksnys (Vilnius University, Technological Sciences, Informatics Engineering – 07T).

The dissertation will be defended at the Council of Scientific Field of Informatics Engineering at Vilnius University, Institute of Mathematics and Informatics:

Chairman

Prof Dr Habil Vytautas Kaminskas (Vytautas Magnus University, Physical Sciences, Informatics – 09 P).

Members:

Prof Dr Habil Gintautas Dzemyda (Vilnius University, Technological Sciences, Informatics Engineering – 07T),

Prof Dr Habil Rimantas Šeinauskas (Kaunas University of Technology, Technological Sciences, Informatics Engineering – 07T),

Prof Dr Habil Kazys Kazlauskas (Vilnius University, Physical Sciences, Informatics – 09 P),

Prof Dr Habil Genadijus Kulvietis (Vilnius Gediminas Technical University, Technological Sciences, Informatics Engineering – 07T).

Opponents:

Prof Dr Dalius Navakauskas (Vilnius Gediminas Technical University, Technological Sciences, Informatics Engineering – 07T),

Prof Dr Dalė Dzemydienė (Mykolas Romeris University, Technological Sciences, Informatics Engineering – 07T).

The dissertation will be defended at the public meeting of the Council of Scientific of Informatics in the auditorium number 203 at Vilnius University, Institute of Mathematics and Informatics, at 1 p. m. on 28th February 2012.

Address: Akademijos st. 4, LT-08663 Vilnius, Lithuania.

The summary of the doctoral dissertation was distributed on 27th of January 2012.

A copy of the doctoral dissertation is available for review at the Library of Vilnius University.

VILNIAUS UNIVERSITETAS

Rasa Lileikytė

ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMAS

Daktaro disertacijos santrauka
Technologijos mokslai, informatikos inžinerija (07T)

Vilnius, 2012

Disertacija rengta 2007–2012 metais Vilniaus universiteto Matematikos ir informatikos institute.

Mokslinis vadovas

prof. habil. dr. Laimutis Telksnys (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – 07T).

Disertacija ginama Vilniaus universiteto Matematikos ir informatikos instituto Informatikos inžinerijos mokslo krypties taryboje:

Pirmininkas

prof. habil. dr. Vytautas Kaminskas (Vytauto Didžiojo universitetas, fiziniai mokslai, informatika – 09 P).

Nariai:

prof. habil. dr. Gintautas Dzemyda (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – 07T),

prof. habil. dr. Rimantas Šeinauskas (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija – 07 T),

prof. habil. dr. Kazys Kazlauskas (Vilniaus universitetas, fiziniai mokslai, informatika – 09 P),

prof. habil. dr. Genadijus Kulvietis (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija – 07 T).

Oponentai:

prof. dr. Dalius Navakauskas (Vilniaus Gedimino technikos universitetas, technologijos mokslai, informatikos inžinerija – 07 T),

prof. dr. Dalė Dzemydienė (Mykolo Romerio universitetas, technologijos mokslai, informatikos inžinerija – 07 T).

Disertacija bus ginama viešame Informatikos inžinerijos mokslo krypties tarybos posėdyje 2012 m. vasario mėn. 28 d. 13 val. Vilniaus universiteto Matematikos ir informatikos instituto 203 auditorijoje.

Adresas: Akademijos g. 4, LT–08663 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2012 m. sausio 27 d.

Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje.

General Characteristic of the Dissertation

Topicality of the problem. Automatic speech recognition technology is widely employed in various areas: customer servicing in large companies as information providing and forwarding telephone calls to reduce waiting time; automated text typing by dictating – no text typing skills are needed and less human resources are used; automated speech translation enables people talk to each other even if they are not capable to speak particular foreign language.

The construction of automatic speech recognition system is a difficult task. The accuracy of speech recognition system depends on characteristics of employed speech recognition features and classifier. Evaluating the accuracy of speech recognition system in ordinary way, the error of speech recognition system has to be calculated for both each explored feature system type and each classifier type. However, this method is limited as error calculation for speech recognition system causes much work and large amount of calculation resources. The amount of this work and calculations can be reduced if the quality of explored feature system is estimated. Accordingly, quality estimation of speech recognition features at the front stage of speech recognition system designing denotes that classification experiments would be not required. Consequently, the new method for quality estimation of speech recognition features is needed that enables to solve speech recognition task in more effective way.

Aim and tasks of the work – to propose a new method for quality estimation of speech recognition features that enables to solve the task of speech recognition in more simple way.

In order to achieve the aim, the following tasks are stated:

1. Propose a new method for quality estimation of speech recognition features that is based on metrics.
2. Propose the set of metrics for quality estimation of speech recognition features.
3. Demonstrate, that the proposed method describes the quality of speech recognition features in Euclidean space and provides the possibility to make the task of speech recognition more simple due to discarding the calculations of classification experiments.

4. Estimate the algorithms complexities of both the proposed method for quality estimation and quality estimation of recognition systems.
5. Develop the experimental base adapted for the experimental researches of quality estimation of speech recognition features.
6. Confirm the correctness of the proposed method for quality estimation of speech recognition features by performing the experimental researches.

Scientific novelty

1. Proposed the method for quality estimation of speech recognition features that is based on metrics.
2. Proposed the set of metrics for quality estimation of speech recognition features.
3. The proposed method describes the quality of speech signals and provides the possibility to make the task of speech recognition more simple, discarding the resources of classification calculations.
4. Demonstrated, that algorithm complexity of the proposed method for quality estimation of speech recognition features is $O(2R\log 2R)$, while algorithm complexity of dynamic time warping recognition system for quality evaluation is $O(R^2)$, where R is vectors number of speech pattern.
5. Developed the experimental base adapted for the experimental researches of quality estimation of speech recognition features.

Methodology of research. We use the knowledge of mathematical statistics, digital signal processing and pattern recognition theory for the theoretical analysis, practical realization and experimental researches. The software was developed using PL/SQL language (Oracle SQL Developer 2.1), C++ language (MinGW 0.2), Matlab R2007b.

Practical value. The proposed method for quality estimation of speech recognition features provides the possibility to perform the quality estimation tasks of speech recognition features in more quality and expedited way. Furthermore, the method provides the possibility to perform the quality estimation tasks of speech recognition systems in more simple and economic way because of discarding the calculations of classification experiments.

Defended propositions

1. Proposed the set of metrics for quality estimation of speech recognition features, consisting of three metrics: feature volume of class boundary, nearest neighbour distances ratio of classes, overstep volume of class boundary. Also it provided the possibility to create useful quality estimation method of speech recognition features.
2. Proposed the new method for quality estimation of speech recognition features that refers to the characteristics of speech recognition systems operating in Euclidean space.
3. Introduced the scale of feature quality index of 0 %–100 %, where 0 % identifies the lowest quality of features and 100 % – the highest quality of features.
4. The proposed method that describes the quality of speech recognition features and provides the possibility to make the recognition task of speech signals more simple due to discarding the calculations of classification experiments.
5. Algorithm complexity of method for quality estimation of speech recognition features is $O(2R\log 2R)$, while algorithm complexity of dynamic time warping recognition system for quality evaluation is $O(R^2)$, where R is vectors number of speech pattern.
6. Developed the experimental base served for experimental researches. It was demonstrated that the new method for quality estimation of speech recognition features enables to describe quality of recognition systems in Euclidean space, using less expenditure.

Approbation of the results. 3 articles focusing on the main dissertation results were published in *ISI Web of Science* journals:

1. Lileikytė, R.; Telksnys, L. 2011. Quality Estimation Methodology of Speech Recognition Features, *Electronics and Electrical Engineering*. 4(110), 113–116. ISSN 1392–1215.

2. Lileikytė, R.; Telksnys, L. 2012. Quality Measurement of Speech Recognition Features in Context of Nearest Neighbour Classifier, *Electronics and Electrical Engineering.* 2(118). ISSN 1392–1215.

3. Lileikytė, R.; Telksnys, L. Quality estimation of speech recognition features for Dynamic Time Warping classifier, *Information Technology and Control.* ISSN 1392–124X. (Accepted for printing).

The main results of the dissertation were reported at 2 scientific conferences and seminars:

1. *Quality Estimation Methodology of Speech Recognition Features* reported in international conference „Electronics“ 2011 m., Kaunas.
2. *Quality Estimation of Speech Recognition Features* reported in international seminar „Institute of Electrical and Electronic Engineers“ 2011 m., Vilnius.

The scope of the scientific work. The scientific work consists of the general characteristic of the dissertation, 3 chapters, conclusions, list of literature, list of publications and annexes. The total scope of the dissertation – 103 pages, 38 pictures, 10 tables, 122 formulas and 2 annexes. Also it was used 118 references.

1. Analysis of quality estimation of speech recognition features

The quality is the measure that describes the adequacy degree of the object’s characteristics with determined requirements. The high object’s quality is achieved if object’s characteristics meet the determined requirements (Bareisa *et al.* 2007; Targamadzė *et al.* 2010; Eidukas *et al.* 2010). The feature system that minimizes classification error is established as quality feature system. In the proposed method the set of metrics is used for the quality estimation of speech recognition features. As well as this, the classification error minimization coincides with the metrics minimization. The classes are well separable if metrics gain minimal values, and classes are overlapped in case of minimal metrics values. Moreover, we introduce feature quality index that summarizes the results of metrics. It’s scale is 0 %–100 %, where 0 % identifies the lowest quality of feature system and 100 % – the highest quality of feature system.

Currently the establishment of quality feature system is based on classification error calculation (Fig. 1) (Webb 2002; Bernado–Mansilla *et al.* 2005; Ho *et al.* 2002; Sotoca *et al.* 2005; Chen *et al.* 2004; Voitovetsky *et al.* 1997; Soryani *et al.* 2005; Gavrilis *et al.* 2008).

During the quality estimation of speech recognition features employing classification error, classification experiments must be performed with each explored feature system. The quality feature system is defined within the lowest classification error. However, the new method is needed as feature quality establishment using classification error has limits. Firstly, the method requires classification experiments performing with each feature system. Let's suppose we have five feature systems under investigation S_1, S_2, S_3, S_4, S_5. In order to choose the quality feature system, the classification process must be run five times (Fig. 1). Furthermore, performing experiments require much work at the learning stage and making models. Equally important, calculation of classification error causes large computational resources, e.g. the usage of testing techniques like leave–one–out significance decrease the performance of classification process (Larranaga *et al.* 2002; Duda *et al.* 2001). Consequently, the new method is needed for quality estimation of speech recognition features that does not require performing classification experiments.

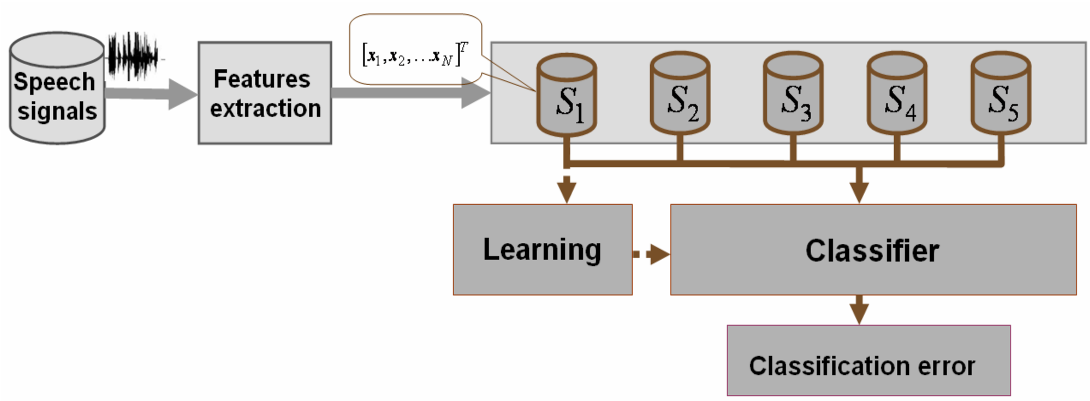

Fig. 1. Features quality estimation using classification error

In order to propose the new method for quality estimation of speech recognition features we reviewed metrics used for tasks as data structure exploration within geometrical metrics (Ho *et al.* 2002; Elizondo *et al.* 2009; Ho 2001; Ho 2002; Ho 2000; Bernado–Mansilla *et al.* 2006; Sotoca *et al.* 2005), synthetic data generation using geometrical metrics (Macia *et al.* 2008). Also for feature selection, extraction employing information theory (Kerroum *et al.* 2008; Arauzo–Azofra *et al.* 2009; Peng *et al.* 2005; Balagani *et al.* 2010; Bao *et al.* 2006; Chen 1971), statistical metrics (Webb 2002;

Theodoridis *et al.* 1999; Costa *et al.* 1999; Malina 1981; Miyamoto *et al.* 2003). The analyzed metrics are presented in Table 1.

Table 1. Groups of metrics and metrics

Group of metric		Metric
Geometrical metrics	Overlap of individual features	Feature efficiency
		Maximal Fisher's discriminant ratio
		Volume of features overlap
	Separability of classes	Features volume of class boundary
		Nearest neighbour distances ratio of classes
		Overstep volume of class boundary
		Cluster number of ε–neighbour
		Error rate of linear classifier
		Error rate of nearest neighbour
		Nonlinearity of linear classifier
		Nonlinearity of nearest neighbour classifier
Information theory metrics		Entropy
		Conditional entropy
		Joint entropy
		Mutual information
Statistical metrics		Hypothesis estimation
		Distance based measures
		Covariance matrices based criterions
		Correlation
		Shape criterions of probability distribution
		Vectors and components of vectors numbers average

2. Quality estimation of speech recognition features

2.1. Proposed method for quality estimation of speech recognition features

Until now the quality of features was estimated by classification error. However, this method is limited due to the need to run classification experiments. Consequently, in this work we propose a new method for quality estimation of speech recognition features within characteristics: it does not require classification experiments performing, also it is suitable for Euclidean space classifiers (classifiers that use Euclidean distance).

The scheme of the proposed method is presented in Fig. 2. Let's suppose that for the given problem we have five feature systems. The inquiry can be stated – how to choose the quality feature system? In the proposed method the quality feature system is determined using the metrics instead of running classification experiments with each feature system. Let's suppose that quality feature system was established S_4 (Fig. 2).

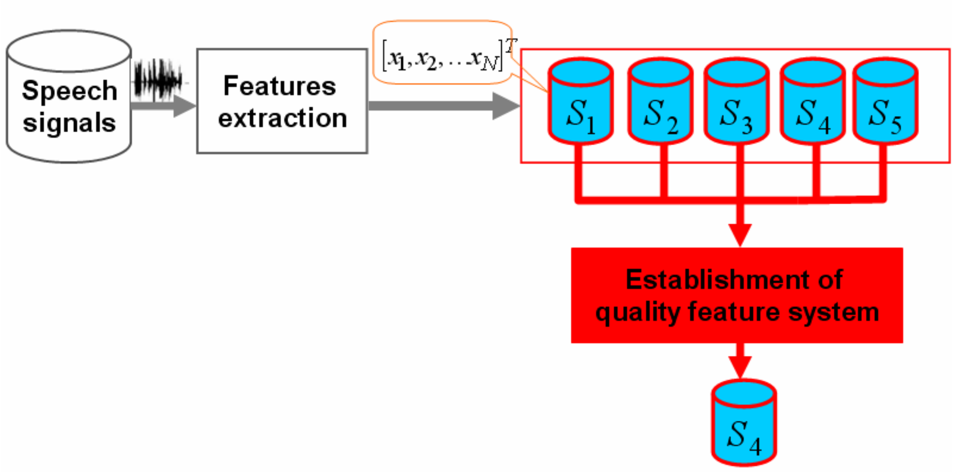

Fig. 2. Features quality estimation using the proposed method

The detailed scheme of the proposed method for quality estimation of speech recognition features is presented in Fig. 3 (Lileikytė *et al*. 2012).

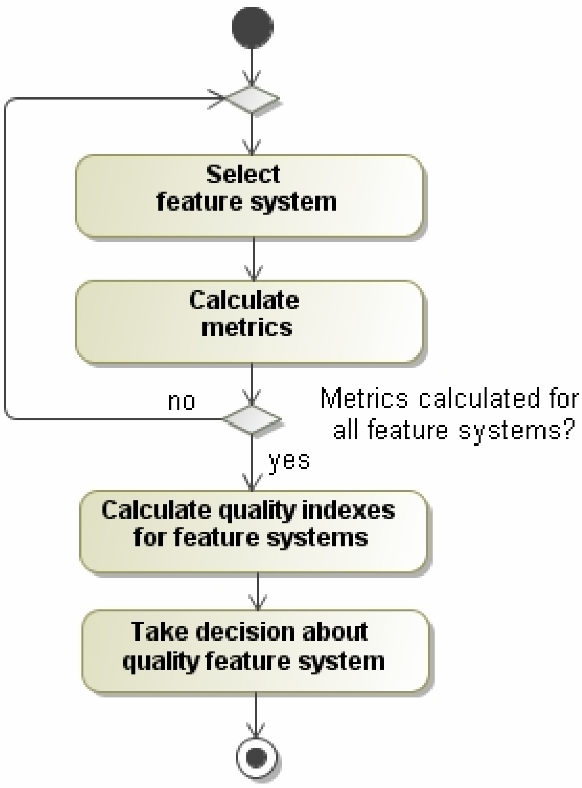

Fig. 3. Detailed scheme of the proposed method for features quality estimation

Furthermore, the proposed method was validated by calculating classification errors for the feature systems. We will describe the stages of the proposed method in detailed. The method is composed from these main stages: selection of feature system, metrics calculation for classes including all explored feature systems, quality indexes calculation for feature systems and finally, decision taking about quality feature system.

Selection of feature system. After feature systems extraction is processed, the metrics are calculated for each feature system.

Metrics calculation. The metrics are calculated for instances combinations that belong to analyzed classes combinations:

$$V_{nm}^{(p,k)}(i,j) = \sigma^{(p,k)}(h_n^i, h_m^j), \tag{1}$$

where $\sigma^{(p,k)}(h_n^i, h_m^j)$ is function of *k*–th metric calculated for h_n^i *n*–th instance of *i*–th class and h_m^j *m*–th instance of *j*–th class, $1 \le k \le K$, K – number of metrics, $1 \le p \le P$ and P is number of feature systems, $1 \le n \le H_i$, H_i is instances number of *i*–th class, $1 \le m \le H_j$, H_j is instances number of *j*–th class, $1 \le i \le C$, $1 \le j \le C$, $i \ne j$, C is number of classes.

The average of metrics is calculated for each combination of classes, including each feature system respectively:

$$VG^{(p,k)}(i,j) = \frac{\sum_{n=1}^{H_i} \sum_{m=1}^{H_j} V_{nm}^{(p,k)}(i,j)}{PBsk}, \tag{2}$$

where $PBsk$ is the number of instances combinations. The low value of metric identifies that classes are well separable.

Quality indexes calculation for feature systems. After the metrics averages are calculated for each classes combinations, the vote is assigned to this feature system that achieved the lowest average value of metric:

$$BG^{(p,k)}(i,j) = f(\arg\min_k VG^{(p,k)}(i,j)), \tag{3}$$

where function $f(\cdot)$ gives 1 if the *k*–th metric achieved the lowest average value for the *p*–th feature system. Otherwise, function gives 0.

Then, the coefficient of votes is calculated for feature system regarding the combination of classes:

$$FG^{p}(i,j)=o(\arg\max_{p}\sum_{k=1}^{K}BG^{(p,k)}(i,j)), \qquad (4)$$

where $o(\cdot)$ function gives 1 if p–th feature system gathered the largest number of votes. In case if more than one feature system got the largest number of votes, the function gives $1/Psk$, where Psk is the number of feature systems that gathered the same largest number of votes. Otherwise, the function gives 0.

Quality index of feature system is calculated as the average of coefficients of votes regarding each feature system:

$$RFG^{p}=\frac{\sum_{i=1}^{C}\sum_{j=1}^{C}FG^{p}(i,j)}{KBsk}\cdot 100, \qquad (5)$$

where $KBsk$ is the number of classes combinations.

Quality index of feature system has 0 %–100 % scale, where 0 % identifies the lowest quality of feature system and 100 % – the highest quality of feature system.

Decision taking about the quality feature system. The quality feature system is determined with the highest quality index:

$$SFG=\arg\max_{p} RFG^{p}. \qquad (6)$$

Validation the adequateness of the proposed method. In order to validate the adequateness of the proposed method classification error has to be calculated for the explored feature systems. The method is adequate if the results of the proposed method coincides with the results while quality establishing by classification error.

Thus, the proposed method for quality estimation of speech recognition features is based on three metrics usage (description of metrics is provided in following section). Algorithm complexity of metric feature volume of class boundary (G1) is $O(2R\log 2R)$, of metric nearest neighbour distances ratio of classes (G2) is $O(2R\log 2R)$ and of metric overstep volume of class boundary (G3) is $O(2R)$, where R – vectors number of speech pattern ($N=2R$, N – vectors number of classes). As a result, the complexity of the proposed method algorithm is $O(2R\log 2R)$. Contrarily, algorithm's complexity of dynamic time warping recognition system is $O(R^2)$.

2.2. Proposed metrics for quality estimation of speech recognition features

The proposed method for quality estimation of speech recognition features is based on the usage of the set of metrics. The proposed set of metrics consists of 3 metrics: features volume of class boundary, nearest neighbour distances ratio of classes, overstep volume of class boundary. The metrics belong to the group of geometrical metrics, due to close relation with classification difficulties as measuring the degree of classes overlapping, boundaries complexity.

Features volume of class boundary (G1). The metric is based on Minimal spanning tree (MJM) (Bock 1971; Kruskal 1956; Prim 1957; Buchin *et al.* 2009; Huang *et al.* 2009; Narasimhan *et al.* 2000) calculation (Ho *et al.* 2002). Let's $L = \{l_1, l_2, \ldots, l_K\}$ is the set of vertexes (vectors) that are connected by edges (Euclidean distances) in MJM and belongs to different classes (Fig. 4). The number of these vectors is normalized by the number of all vectors:

$$G1 = \frac{K}{N}, \tag{7}$$

where K – number of vectors connected by edges and belonging to different classes, N is number of all vectors.

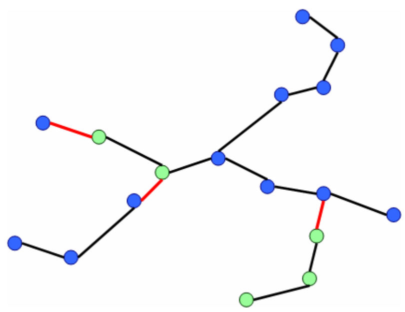

Fig. 4. Features amount of the classes boundaries

The minimal metric's value is the positive value near zero and shows that classes are well separated. The maximal metric's value is one and shows that classes are overlapped. The complexity of metric's algorithm is $O(N \log N)$.

Nearest neighbour distances ratio of classes (G2). The metric is calculated as distances to the nearest vector of the same class and to the nearest vector of the opposite class (Bernado–Mansilla *et al.* 2005). First, for each vector Euclidean distance is calculated both to the nearest vector of the same class and of the opposite class (Fig. 5). Then, the ratio is calculated of these distances:

$$G2 = \frac{\sum_{i=1}^{C}\sum_{n=1}^{N_i}\min_k d(x_n^i, x_k^i)}{\sum_{i=1, j=1, i\neq j}^{`C}\sum_{n=1}^{N_i}\min_m d(x_n^i, x_m^j)}, \quad (8)$$

where $\min_k d(x_n^i, x_k^i)$ is minimal Euclidean distance between x_n^i – n–th vector of i–th class and x_k^i – k–th vector of i–th class, $1 \le n \le N_i$, $1 \le k \le N_i$, N_i is vectors number of i–th class, $1 \le i \le C$, $1 \le j \le C$, $i \neq j$, C – number of classes, $\min_m d(x_n^i, x_m^j)$ is minimal Euclidean distance between x_n^i – n–th vector of i–th class and x_m^j – m–th vector of j–th class, $1 \le m \le N_j$, N_j is vectors number of j–th class.

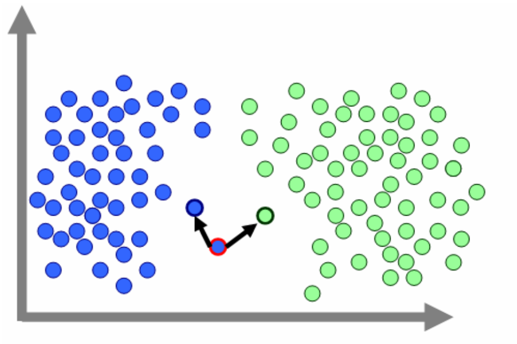

Fig. 5. Distances to the nearest vectors of the same and opposite classes

The minimal metric's value is zero and shows that classes are well separated. In case the distance within the same class is smaller than between the opposite classes, the largest metric's value is one and shows that classes are overlapped. Contrarily, the maximal metric's value is bigger than one and not defined. The complexity of metric's algorithm is $O(N \log N)$.

Overstep volume of class boundary (G3). Let us suppose that every class is represented by the circle (Li *et al.* 1998). The radius of the circle is defined as the distance from the centre to the farthest vector of the class, where the centre is the mean of the class. The radius is calculated:

$$r^i = \max_n d(\mu^i, x_n^i), \quad (9)$$

where $\max_n d(\mu^i, x_n^i)$ is maximal Euclidean distance between μ^i – centre of i–th class and x_n^i – n–th vector of i–th class, $1 \le i \le C$, $1 \le n \le N_i$.

The number of vectors is found that occur in the area of opposite class (Fig. 6). Then this number is normalized by product of classes $(C-1)$ and vectors number:

$$G3 = F[d(\mu^i, x_m^j) \le r^i] \cdot \frac{1}{C-1} \cdot \frac{1}{N}, \tag{10}$$

where $d(\mu^i, x_m^j)$ – Euclidean distance between the centre of *i*–th class and *m*–th vector of *j*–th class, $1 \le m \le N_j$, N – number of all vectors. Function $F[\cdot]$ returns the number of distances that satisfy the condition of class boundary overstep volume of $d(\mu^i, x_m^j) \le r^i$.

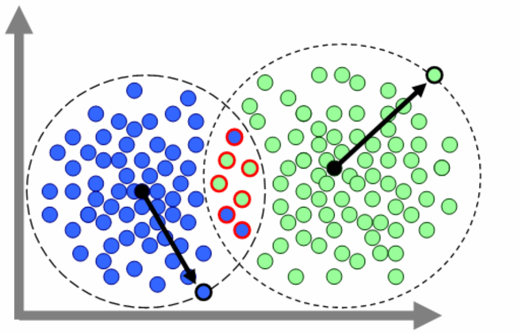

Fig. 6. Vectors of classes falling into circle of other class

The minimal metric's value is zero and shows that classes are well separated. The maximal metric's value is one and shows that classes are overlapped. The complexity of metric's algorithm is $O(N)$.

3. Experimental researches for quality estimation of speech recognition features

Computer aided experimental researches were performed to validate the adequateness of the proposed method.

3.1. Data and conditions of the experimental researches

The experimental researches were performed with two feature systems: 12th order perceptual linear prediction cepstrum coefficients (TSPKK) (Hermansky *et al.* 1985; Rabiner *et al.* 1993), 12th order real–cepstrum coefficients (KK) (Furui 2001).

For the experiment we used the proposed set of metrics: feature volume of class boundary (G1), nearest neighbour distances ratio of classes (G2), overstep volume of class boundary (G3). We employed 14 different non – contextual phonemes representing different classes: [i], [a], [e], [o:], [s], [k], [t], [t'], [m], [n'], [r], [r'], [j], [s'] (where „ ' " means soft consonant, „ : " – long vowel). These phonemes are commonly used in

Lithuanian language (Raškinis *et al.* 2009). The phonemes are extracted from short connected phrases corpus VDU–TRI4 of Vytautas Magnus University (Raškinis *et al.* 2004). Experimental researches were performed with 3 data sets: DM1 with speaker male, DM2 with speaker female, DM3 with four speakers – two males and two females. Each data set was composed from 14 classes, each class having 100 instances. Calculations were made for 91 pairs of classes, including 10 000 instances combinations for each pair of classes. Two classifiers were employed in order to confirm the correctness of the proposed method: nearest neighbour (AK) classifier (Cover *et al.* 1967; Golipour *et al.* 2009; Hua *et al.* 2010) and dynamic time warping (DLSK) classifier (Sakoe *et al.* 1978; Rabiner *et al.* 1978; Bin Amin *et al.* 2008). Both classifiers used Euclidean distance. 70 % of data was used for training and 30 % for testing. Algorithm complexity of AK recognition system is $O(2R)$, where R – the number of vectors of speech pattern ($N = 2R$, where N is vectors number of classes). Therefore, algorithm complexity of DLSK recognition system is $O(R^2)$.

3.2. Experimental researches for quality estimation of features

In this section the experimental results of the proposed method for quality estimation of speech recognition features are presented by providing quality indexes of feature systems. In order to validate the adequateness of proposed method, classification errors were calculated respectively to AK and DLSK classifiers with 80 % confidence level.

The results of experimental researches of *DM1 data set*. The results of quality indexes of TSPKK and KK feature systems are presented in Fig. 7.

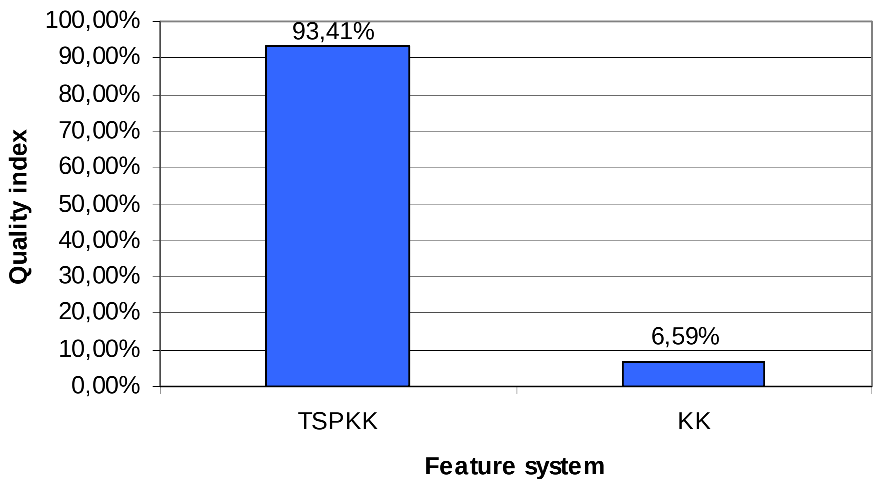

Fig. 7. Quality index of feature system (DM1 data set)

Consequently, observing the results it can be stated that quality index of TSPKK feature system is higher comparing to KK feature system. Therefore, using the proposed method for feature quality estimation, TSPKK feature system was established quality feature system.

To validate the adequateness of the proposed method, both AK and DLSK classification errors were calculated for TSPKK and KK feature systems. The results of classification errors are displayed in Table 2.

Table 2. Classification error (DM1 data set)

Classifier	System	Classification error
AK	TSPKK	$9{,}14 \pm 1{,}11$ %
	KK	$13{,}48 \pm 1{,}51$ %
DLSK	TSPKK	$3{,}63 \pm 0{,}76$ %
	KK	$6{,}17 \pm 1{,}06$ %

The results showed that the lowest AK classification error was achieved for TSPKK feature system. As well as this, the lowest DLSK classification error was achieved for TSPKK feature system. Therefore, employing both AK and DLSK classifiers it was established that TSPKK is quality feature system.

The results of experimental researches of *DM2 data set*. The results of estimating quality indexes of TSPKK and KK feature systems are displayed in Fig. 8.

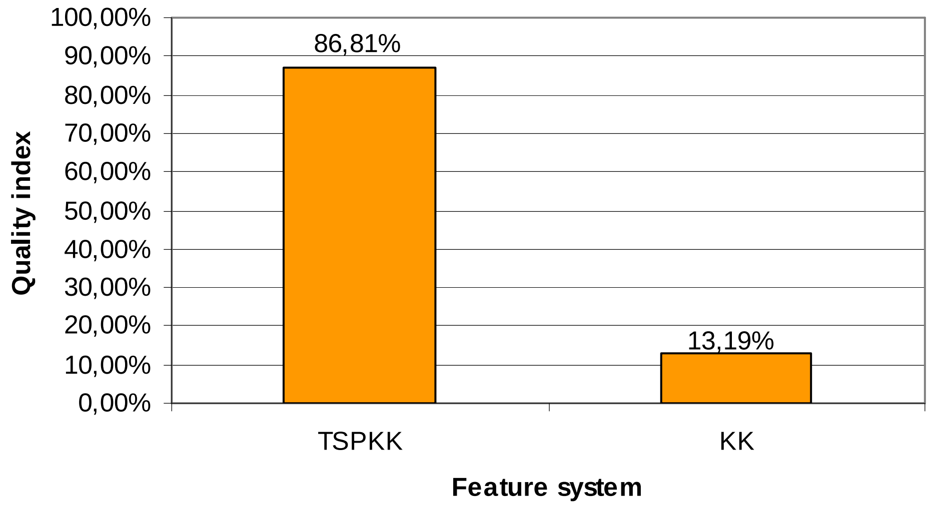

Fig. 8. Quality index of feature system (DM2 data set)

The experimental results showed that quality index of TSPKK feature system was higher than quality index of KK feature system. Therefore, employing the proposed method for feature quality estimation, it was established that TSPKK was quality feature system.

As well as this, AK and DLSK classification errors were calculated for TSPKK and KK feature systems. The results of classification errors are displayed in Table 3.

Table 3. Classification error (DM2 data set)

Classifier	System	Classification error
AK	TSPKK	$8{,}12 \pm 0{,}93$ %
	KK	$13{,}12 \pm 1{,}50$ %
DLSK	TSPKK	$2{,}48 \pm 0{,}62$ %
	KK	$4{,}67 \pm 0{,}73$ %

Accordingly, the results showed that the lowest AK classification error was achieved for TSPKK feature system. In addition to this, the lowest DLSK classification error was gained for TSPKK feature system. Consequently, TSPKK was established the quality feature system.

The results of experimental researches of *DM3 data set*. The results of quality indexes of TSPKK and KK feature systems are presented in Fig. 9.

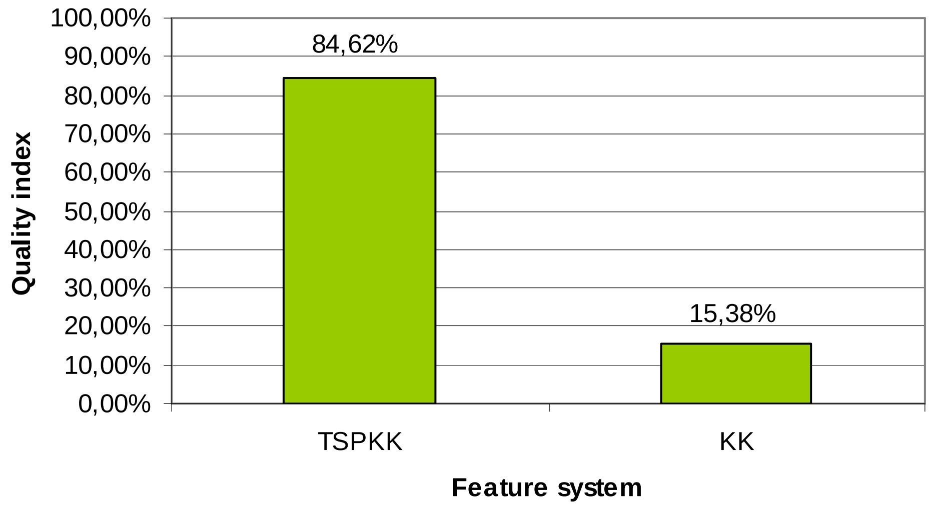

Fig. 9. Quality index of feature system (DM3 data set)

The results defined that TSPKK feature system was established quality feature system, due to quality index of TSPKK feature was higher than quality index of KK feature system.

Equally important, AK and DLSK classification errors were calculated for TSPKK and KK feature systems. The results of classification errors are provided in Table 4.

Table 4. Classification error (DM3 data set)

Classifier	System	Classification error
AK	TSPKK	$10{,}09 \pm 1{,}18$ %
	KK	$13{,}66 \pm 1{,}55$ %
DLSK	TSPKK	$3{,}53 \pm 0{,}63$ %
	KK	$5{,}70 \pm 0{,}78$ %

Consequently, the experimental results showed that the lowest AK classification error was calculated for TSPKK feature system. Therefore, the lowest DLSK classification error was achieved for TSPKK feature system. Accordingly, TSPKK was established the quality feature system.

In conclusion, the experimental results of features quality estimation of speech recognition features employing the proposed method coincided with the results of quality estimation using the classification error. In both cases, with all three data sets, it was established that TSPKK was the quality feature system. Therefore, the highest quality index identified the feature system with the lowest classification error. Consequently, the results of the experimental researches confirmed the correctness of the proposed method. As a result, it was demonstrated that quality of speech recognition features in Euclidean space describes the quality of recognition system and doesn't require performing classification experiments.

General Conclusions

After developing the proposed method for quality estimation of speech recognition features, as well as after performing the experimental researches according this method, there were formulated the following conclusions:

1. Proposed the method for quality estimation of speech recognition features that is based on metrics.
2. Proposed the set of metrics for quality estimation of speech recognition features, consisting of three metrics: feature volume of class boundary, nearest neighbour distances ratio of classes, overstep volume of class boundary.
3. Demonstrated, that quality of speech recognition features in Euclidean space describes the quality of recognition system and doesn't require performing classification experiments.
4. Introduced the scale of feature quality index of 0 %–100 %, where 0 % identifies the lowest quality of feature system and 100 % – the highest quality of feature system.
5. Demonstrated, that algorithm complexity of method for quality estimation of speech recognition features is $O(2R \log 2R)$, while algorithm complexity of dynamic time warping recognition system is $O(R^2)$, where R is vectors number of speech pattern.
6. Developed the experimental base adapted for the experimental researches of quality estimation of speech recognition features.
7. Confirmed the correctness of the proposed method for quality estimation of speech recognition features in case of nearest neighbour and dynamic time warping recognition systems.
8. Demonstrated, that in case of the first data set (speaker – male) when the quality index of perceptual linear prediction cepstrum coefficients gained 93,41 % and cepstrum coefficients gained 6,59 %, it was achieved the following speech recognition results:
    1) Recognition system, that used perceptual linear prediction cepstrum coefficients and nearest neighbour classifier gained $9{,}14 \pm 1{,}11$ % classification error. Moreover, recognition system, that used cepstrum coefficients and nearest neighbour classifier achieved $13{,}48 \pm 1{,}51$ % classification error.
    2) Recognition system, that used perceptual linear prediction cepstrum coefficients and dynamic time warping classifier gained $3{,}63 \pm 0{,}76$ %

classification error. Further, recognition system, that used cepstrum coefficients and dynamic time warping classifier achieved $6{,}17 \pm 1{,}06$ % classification error.

9. Demonstrated, that in case of the second data set (speaker – female) when the quality index of perceptual linear prediction cepstrum coefficients gained 86,81 % and cepstrum coefficients gained 13,19 %, it was achieved the following speech recognition results:
    1) Recognition system, that used perceptual linear prediction cepstrum coefficients and nearest neighbour classifier gained $8{,}12 \pm 0{,}93$ % classification error. Therefore, recognition system, that used cepstrum coefficients and nearest neighbour classifier achieved $13{,}12 \pm 1{,}50$ % classification error.
    2) Recognition system, that used perceptual linear prediction cepstrum coefficients and dynamic time warping classifier gained $2{,}48 \pm 0{,}62$ % classification error. Besides, recognition system, that used cepstrum coefficients and dynamic time warping classifier achieved $4{,}67 \pm 0{,}73$ % classification error.
10. Demonstrated, that in case of the third data set (two speakers males and two females) when the quality index of perceptual linear prediction cepstrum coefficients gained 84,62 % and cepstrum coefficients gained 15,38 %, it was achieved the following speech recognition results:
    1) Recognition system, that used perceptual linear prediction cepstrum coefficients and nearest neighbour classifier gained $10{,}09 \pm 1{,}18$ % classification error. Therefore, recognition system, that used cepstrum coefficients and nearest neighbour classifier achieved $13{,}66 \pm 1{,}55$ % classification error.
    2) Recognition system, that used perceptual linear prediction cepstrum coefficients and dynamic time warping classifier gained $3{,}53 \pm 0{,}63$ % classification error. Thus, recognition system, that used cepstrum coefficients and dynamic time warping classifier achieved $5{,}70 \pm 0{,}78$ % classification error.

11. Purposeful to continue researches of quality estimation of recognition features dealing with recognition tasks of mechanical and biological dynamic's states systems.

List of Published Works on the Topic of the Dissertation

In the reviewed scientific periodical publications

Lileikytė, R.; Telksnys, L. 2011. Quality Estimation Methodology of Speech Recognition Features, *Electronics and Electrical Engineering* 4 (110), 113–116. ISSN 1392–1215 (ISI Web of Science Journal List).

Lileikytė, R.; Telksnys, L. 2012. Quality Measurement of Speech Recognition Features in Context of Nearest Neighbour Classifier, *Electronics and Electrical Engineering* 2(118). ISSN 1392–1215 (ISI Web of Science Journal List).

Lileikytė, R.; Telksnys, L. Quality estimation of speech recognition features for Dynamic Time Warping classifier, *Information Technology and Control*. ISSN 1392–124X. (Accepted for printing). (ISI Web of Science Journal List).

About the author

Rasa Lileikytė was born in Telšiai, on 30 of November 1982.

Bachelor degree in Applied Informatics, Faculty of Informatics, Vytautas Magnus University, 2005. Master of Applied Informatics, Faculty of Informatics, Vytautas Magnus University, 2007. In 2007–2012 – PhD student of Vilnius University, Institute of Mathematics and Informatics.

ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMAS

Mokslo problemos aktualumas. Automatinis šnekos atpažinimas yra taikomas įvairiose žmonių veiklose: klientų aptarnavimas didelėse kompanijose – informacijos teikimas, skambučių priėmimų centrų automatizavimas sumažinant laukimo laiką; automatizuoto teksto rinkimas diktuojant – teksto surinkimas nereikalauja surinkimo įgūdžių ir sunaudojama mažiau žmonių darbo resursų; pasinaudodami automatiniu kalbos vertimu, žmonės gali susikalbėti ir nemokėdami reikiamos kalbos.

Šnekos signalų atpažinimo sistemų sudarymas yra sudėtingas uždavinys. Šnekos signalų atpažinimo sistemų tikslumas priklauso nuo šnekos signalus aprašančių požymių ir šiuos požymius naudojančių klasifikatorių savybių. Tradiciškai vertinant atpažinimo sistemų tikslumą, kiekvienai pasirinktai požymių sistemai ir kiekvienam klasifikatoriaus tipui tenka atlikti atpažinimo tikslumo skaičiavimus. Tačiau šis metodas yra neefektyvus, nes atpažinimo sistemos klaidos apskaičiavimas reikalauja daug darbo ir didelių skaičiavimo resursų. Tokių darbų apimtis galima sumažinti – atpažinimo sistemos konstravimo pradiniame etape atlikus pasirenkamų požymių kokybės vertinimą. Taigi, reikalingas naujas šnekos signalų atpažinimo požymių kokybės vertinimo metodas, kuris palengvintų šnekos signalų atpažinimo uždavinį.

Darbo tikslas ir uždaviniai. Pagrindinis šio darbo tikslas – pateikti šnekos signalų atpažinimo požymių kokybės vertinimo metodą, suteikiantį galimybę palengvinti šnekos signalų atpažinimo uždavinį.

Darbo tikslui pasiekti reikia išspręsti šiuos uždavinius:

1. Pateikti naują šnekos signalų atpažinimo požymių kokybės vertinimo metodą, grindžiamą metrikų naudojimu.
2. Pateikti metrikų rinkinį šnekos signalų atpažinimo požymių kokybei vertinti.
3. Parodyti, kad šnekos signalų požymių kokybę aprašantis metodas Euklido erdvėje suteikia galimybę supaprastinti šnekos signalų atpažinimo uždavinį, atsisakant klasifikatoriaus įtakos skaičiavimų.
4. Įvertinti šnekos signalų požymių kokybės ir atpažinimo sistemų kokybės vertinimo metodų algoritmų sudėtingumą.
5. Sukurti šnekos signalų atpažinimo požymių kokybės vertinimo tyrimams pritaikytą eksperimentinę bazę.
6. Patvirtinti eksperimentiškai pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumą.

Mokslinis naujumas

1. Pateiktas šnekos signalų atpažinimo požymių kokybės vertinimo metodas, grindžiamas metrikų naudojimu.
2. Pateiktas metrikų rinkinys šnekos signalų atpažinimo požymių kokybei vertinti.
3. Pasiūlytas šnekos signalų požymių kokybę aprašantis metodas, suteikiantis galimybę supaprastinti šnekos signalų atpažinimo uždavinį, atsisakant klasifikatoriaus įtakos skaičiavimų.
4. Parodyta, kad šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R \log 2R)$, o dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$, R – šnekos signalo objekto vektorių skaičius.
5. Sukurta eksperimentinė bazė, leidžianti atlikti eksperimentinius tyrimus šnekos signalų požymių ir šnekos signalų atpažinimo sistemų kokybei vertinti.

Tyrimų metodika. Teorinei analizei, praktinei realizacijai ir tyrimams panaudotos matematinės statistikos, skaitmeninio signalų apdorojimo bei atpažinimo teorijos žinios. Programinė įranga sukurta naudojant PL/SQL kalbą (Oracle SQL Developer 2.1), C++ kalbą (MinGW 0.2) ir Matlab R2007b integruotą programų kūrimo aplinką.

Praktinė vertė. Pasiūlytas šnekos signalų atpažinimo požymių kokybės vertinimo metodas suteikia galimybę kokybiškiau atlikti ir paspartinti šnekos signalų atpažinimo požymių kokybės vertinimo darbus, o taip pat supaprastinti, taupiau spręsti šnekos signalų atpažinimo sistemų kokybės vertinimo uždavinius, atsisakant klasifikatoriaus įtakos skaičiavimų.

Ginamieji teiginiai

1. Pateiktas metrikų rinkinys šnekos signalų atpažinimo požymių kokybei vertinti, susidedantis iš trijų metrikų – požymių kiekio klasės ribose, klasių artimiausių kaimynų atstumų santykio, klasės ribos peržengimo kiekio – suteikė galimybę sukurti naudingą šnekos signalų atpažinimo požymių kokybės vertinimo metodą.

2. Pateiktas naujas šnekos signalų atpažinimo požymių kokybės vertinimo metodas atspindi šnekos atpažinimo sistemų, veikiančių Euklido erdvėje, savybes.
3. Įvesta požymių kokybės rodiklių skalė 0 %–100 %, kur 0 % rodo žemiausią požymių sistemos kokybę, o 100 % – aukščiausią požymių sistemos kokybę, leidžia matuoti šnekos signalų atpažinimo požymių vertingumą.
4. Pasiūlytas šnekos signalų požymių kokybę aprašantis metodas suteikia galimybę supaprastinti šnekos signalų atpažinimo uždavinį, atsisakant klasifikatoriaus įtakos skaičiavimų.
5. Šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R \log 2R)$, o dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$, R – šnekos signalo objekto vektorių skaičius.
6. Sukurta eksperimentinė bazė padėjo eksperimentiškai parodyti, kad naujasis šnekos signalų atpažinimo požymių kokybės vertinimo metodas Euklido erdvėje leidžia mažesnėmis sąnaudomis aprašyti atpažinimo sistemų kokybę.

Darbo rezultatų aprobavimas

Disertacijos tema atspausdinti 3 moksliniai straipsniai žurnaluose, įtrauktuose į ISI Web of Science sąrašą:

1. Lileikytė, R.; Telksnys, L. 2011. Quality Estimation Methodology of Speech Recognition Features, *Electronics and Electrical Engineering.* 4(110), 113–116. ISSN 1392–1215.

2. Lileikytė, R.; Telksnys, L. 2012. Quality Measurement of Speech Recognition Features in Context of Nearest Neighbour Classifier, *Electronics and Electrical Engineering.* 2(118). ISSN 1392–1215.

3. Lileikytė, R.; Telksnys, L. 2012. Quality estimation of speech recognition features for Dynamic Time Warping classifier, *Information Technology and Control.* ISSN 1392–124X. (Priimta spausdinimui).

Disertacijoje atliktų tyrimų rezultatai paskelbti 2 mokslinėse konferencijose ir seminaruose:

1. *Quality Estimation Methodology of Speech Recognition Features* pristatytas tarptautinėje konferencijoje „Elektronika“ 2011 m., Kaune.
2. *Šnekos atpažinimo požymių kokybės vertinimas* pristatytas tarptautiniame seminare „Institute of Electrical and Electronic Engineers“ 2011 m., Vilniuje.

Darbo apimtis. Darbą sudaro bendra darbo charakteristika, 3 skyriai, išvados, literatūros sąrašas, publikacijų sąrašas ir priedai. Bendra disertacijos apimtis – 103 puslapiai, neįskaitant priedų, 38 iliustracijos, 10 lentelių, 122 formulės ir du priedai. Rašant disertaciją buvo panaudoti 118 literatūros šaltinių.

Pirmasis disertacijos skyrius skirtas šnekos atpažinimo požymių kokybės vertinimo analitinei apžvalgai. Apžvelgtos šnekos signalų atpažinimo sistemą sudarančios pagrindinės dalys. Atlikta šnekos signalų atpažinimo požymių kokybės vertinimo metodų bei metrikų analitinė apžvalga. Skyriaus pabaigoje formuluojamos išvados ir tikslinami disertacijos uždaviniai.

Antrajame skyriuje pateiktas sukurtas metodas šnekos signalų atpažinimo požymių kokybei vertinti. Metodas yra grindžiamas metrikų naudojimu, pateiktos metrikų charakteristikos. Taip pat pateiktas sukurtos programinės įrangos aprašymas.

Trečiajame skyriuje pateikti eksperimentinių tyrimų rezultatai, kurių tikslas yra eksperimentiškai patvirtinti pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumą. Gauti rezultatai eksperimentiškai patvirtino metodo teisingumą.

Bendrosios išvados

Pasiūlius metodą šnekos atpažinimo požymių kokybei vertinti, pagal šį metodą atlikus eksperimentus, suformuluotos išvados:

1. Pateiktas šnekos signalų atpažinimo požymių kokybės vertinimo metodas, grindžiamas metrikų naudojimu.
2. Pateiktas metrikų rinkinys šnekos signalų atpažinimo požymių kokybei vertinti, susidedantis iš trijų metrikų: požymių kiekio klasės ribose, klasių artimiausių kaimynų atstumų santykio, klasės ribos peržengimo kiekio.

3. Parodyta, kad šnekos signalų atpažinimo požymių kokybė Euklido erdvėje aprašo atpažinimo sistemų kokybę ir nereikalauja klasifikavimo eksperimentų vykdymo.
4. Įvesta požymių kokybės rodiklių skalė 0 %–100 %, kur 0 % rodo žemiausią požymių sistemos kokybę, o 100 % – aukščiausią požymių sistemos kokybę.
5. Parodyta, kad šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R\log 2R)$, kai dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$, kur R – šnekos signalo objekto vektorių skaičius.
6. Sukurta šnekos signalų atpažinimo požymių kokybės vertinimo tyrimams pritaikyta eksperimentinė bazė.
7. Patvirtintas pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumas artimiausio kaimyno ir dinaminio laiko skalės atpažinimo sistemų pavyzdžiais.
8. Parodyta, kad pirmojo duomenų rinkinio atveju (kalbėtojas vyras), kai tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 93,41 %, o kepstro koeficientų požymių sistemos kokybės rodiklis buvo 6,59 %, signalų atpažinimo rezultatai buvo tokie:
    - Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $9{,}14 \pm 1{,}11$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $13{,}48 \pm 1{,}51$ % klaidų.
    - Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $3{,}63 \pm 0{,}76$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $6{,}17 \pm 1{,}06$ % klaidų.
9. Parodyta, kad antrojo duomenų rinkinio atveju (kalbėtoja moteris), kai tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės

rodiklis buvo 86,81 %, o kepstro koeficientų požymių sistemos kokybės rodiklis buvo 13,19 %, signalų atpažinimo rezultatai buvo tokie:

- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $8{,}12 \pm 0{,}93$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $13{,}12 \pm 1{,}50$ % klaidų.
- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $2{,}48 \pm 0{,}62$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $4{,}67 \pm 0{,}73$ % klaidų.

10. Parodyta, kad trečiojo duomenų rinkinio atveju (du kalbėtojai vyrai, dvi kalbėtojos moterys), kai tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 84,62 %, o kepstro koeficientų požymių sistemos kokybės rodiklis buvo 15,38 %, signalų atpažinimo rezultatai buvo tokie:
    - Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $10{,}09 \pm 1{,}18$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $13{,}66 \pm 1{,}55$ % klaidų.
    - Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $3{,}53 \pm 0{,}63$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $5{,}70 \pm 0{,}78$ % klaidų.
11. Tikslinga tęsti atpažinimo požymių kokybės vertinimo metodų tyrimus sprendžiant mechaninių ir biologinių dinaminių sistemų būsenų atpažinimo problemas.

Trumpos žinios apie autorių

Rasa Lileikytė gimė 1982 m. lapkričio 30 d. Telšiuose.

2005 m. įgijo taikomosios informatikos bakalauro laipsnį Vytauto Didžiojo universiteto Informatikos fakultete. 2007 m. įgijo taikomosios informatikos mokslo magistro laipsnį Vytauto Didžiojo universiteto Informatikos fakultete. 2007–2012 m. Vilniaus universiteto Matematikos ir informatikos instituto doktorantė.

References

Arauzo–Azofra, A.; Aznarte, J. L. M.; Benitez, J. M. 2009. Empirical Study of Individual Feature Evaluators and Cutting Criteria for Feature Selection in Classification, *Ninth International Conference on Intelligent Systems Design and Applications*, 541–546.

Balagani, K. S.; Phoha, V. V. 2010. On the Feature Selection Criterion Based on an Approximation of Multidimensional Mutual Information, *IEEE Transactions on pattern analysis and machine intelligence* 32(7): 1342–1343.

Bao, M.; Guan, L.; Li, X.; Tian, J.; Yang, J. 2006. Feature Extraction Using Histogram Entropies of Euclidean Distances for Vehicle Classification, *International Conference on Computational Intelligence and Security* V.1: 668–673.

Bareisa, E.; Jusas, V.; Motiejunas, K.; Seinauskas, R. 2007. The Criteria of Functional Delay Test Quality Assessment, *IEEE 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools (DSD 2007)*, 207–214.

Bernado–Mansilla, E.; Ho, T. K. 2005. Domain of competence of XCS classifier system in complexity measurement space, *IEEE Transactions on Evolutionary Computation* 9(1): 82–104.

Bernado–Mansilla, E.; Ho, T. K. and Orriols A. 2006. Data Complexity and Evolutionary Learning. *Data Complexity in Pattern Recognition, Springer*, 115–134.

Bin Amin, T.; Mahmood, I. 2008. Speech Recognition using Dynamic Time Warping, *2nd International Conference on Advances in Space Technologies*, 74–79.

Bock, F. 1971. An algorithm to construct a minimum direct spanning tree in a direct network, *Developments in Operations Research, Gordon and Breach, New York*, 29-44.

Buchin, K.; Mulzer, W. 2009. Delaunay triangulations in O(sort(n)) time and more, *Proc. 50th IEEE Symposium on Foundations of Computer Science*, 139–148.

Chen, C. H. 1971. Theoretical Comparison of a Class of Feature Selection Criteria in Pattern Recognition, *IEEE transaction on computers* C-20(9): 1054–1056.

Chen, X.; Tian, J.; Cheng, J.; Yang, X. 2004. Segmentation of Fingerprint Images Using Linear Classifier, *EURASIP Journal on Applied Signal Processing, EURASIP JASP* 4: 480–494.

Costa, P.; Grouffaud, J.; Larzabal, P.; Clergeot, H.; 1999. Estimation of the number of signals from features of the covariance matrix: a supervised approach, *IEEE Transactions on Signal Processing* 47(11): 3108–3115.

Cover, T. M. and Hart, P. E. 1967. Nearest neighbour pattern classification, *IEEE Transactions on Information Theory* IT-13(1): 21–27.

Duda, R. O.; Hart, P. E and Stork, D. G. 2001. *Pattern classification, Second edition*. John Wiley. 654 p.

Eidukas, D., Kalnius, R. 2010. Models Quality of Electronics Products, *Electronics and Electrical Engineering* 2(98): 29–34.

Elizondo, D. A.; Birkenhead, R.; Gamez, M.; Garcia, N.; Alfaro, E. 2009. Estimation of Classification Complexity. *Proceedings of International Joint Conference on Neural Networks*, 764–770.

Furui, S. 2001. *Digital Speech Processing, Synthesis, and Recognition*. Second edition, Revised and Expanded. Signal Processing and Communications. 452 p.

Gavrilis, D.; Tsoulos, I. G.; Dermatas, E. 2008. Selecting and constructing features using grammatical evolution, *Pattern Recognition Letters* 29(9): 1358–1365.

Golipour, L.; O'Shaughnessy, D. 2009. Context–independent phoneme recognition using a K-Nearest Neighbour classification approach, *IEEE International Conference on Acoustics Speech and Signal Processing*, 1341–1344.

Hermansky, H.; Hanson, B. A.; Wakita, H. 1985. Perceptually based linear predictive analysis of speech, Acoustics, Speech, and Signal Processing, *IEEE International Conference on ICASSP '85* V.10: 509–512.

Ho, T. K. 2001. Data complexity analysis for classifier combination, *Springer Berlin, Heidelberg, Multiple classifier systems* V. 2096: 53–67.

Ho, T. K. 2000. Complexity of classification problems and comparative advantages of combined classifiers, *Proceedings of the 1 st International Workshop Multiple Classifier Systems*, 97–106.

Ho, T. K. 2002. A data complexity analysis of comparative advantages of decision forest constructors, *Pattern Analysis and Applications* V. 5: 102–112.

Ho, T. K.; Basu, M. 2002. Complexity measures of supervised classification problems, *IEEE transactions on pattern analysis and machine intelligence* 24 (3): 289–300.

Hua, O.; Ji, A.; He, O. 2010. Multiple Real–valued K Nearest Neighbor Classifiers System by Feature Grouping, *IEEE International Conference on Systems Man and Cybernetic (SMC)*, 3922–3925.

Huang, F.; Gao, G.; Wang, Y. 2009. Comparison of Prim and Kruskal on Shanghai and Shenzhen 300 Index hierarchical structure tree, *International Conference on Web Information Systems and Mining*, 237–241.

Kerroum, M. A.; Hammouch, A.; Aboutajdine, D.; Bellaachia, A. 2008. Using the maximum Mutual Information criterion to textural Feature Selection for satellite image classification. *IEEE Symposium on Computers and Communications*. ISCC 2008, 1005–1009.

Kruskal, J. B. 1956. On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem, *In Proceedings of the American Mathematical Society* 7(1): 48–50.

Larranaga, P.; Lozano, J. A. 2002. *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation (Genetic Algorithms and Evolutionary Computation).* Kluwer Academic Publishers. 382 p.

Lileikytė, R.; Telksnys, L. 2012. Quality Measurement of Speech Recognition Features in Context of Nearest Neighbour Classifier, *Electronics and Electrical Engineering* 2(118). ISSN 1392–1215.

Li, Z.; Ditang, F. 1998. The research on speech feature representation method and distance measure method, *Pattern recognition, 9th International Conference* V. 1: 631–633.

Macia, N.; Bernado-Mansilla, E.; Orriols-Puig, A. 2008. Preliminary approach on synthetic data sets generalization based on class separability measure, *IEEE International Conference on Pattern Recognition*, 1–4.

Malina, W. 1981. On an Extended Fisher Criterion for Feature Selection, *IEEE Transactions on pattern analysis and machine intelligence* 3(5): 611–614.

Miyamoto, T.; Uchimura, S.; Hamamoto, Y.; Iizuka, N.; Oka, M.; Yamada-Okabe, H. 2003. Comparative study of feature selection methods on microarray data, *IEEE EMBS Asian–Pacific Conference on Biomedical Engineering*, 82–83.

Narasimhan, G.; Zachariasen, M. and Zhu, J. 2000. Experiments with computing geometric minimum spanning trees, In Proceedings of ALENEX'00, 183–196.

Peng, H. C.; Long, F. and Ding, C. 2005. Feature selection based on mutual information: criteria of max–dependency, max–relevance, and min–redundancy, *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(8): 1226–1238.

Prim, R. C. 1957. Shortest connection networks and some generalizations, *In Bell System Technical Journal* 36: 1389–1401.

Rabiner, L. R.; Juang, B. H. 1993. *Fundamentals of speech recognition*. Prentice-Hall, New Jersey. 507 p.

Rabiner, L. R.; Rosenberg, A. E.; Levinson, S.E. 1978. Considerations in dynamic time warping algorithms for discrete word recognition, *IEEE transactions on Acoustics, Speech and Signal Processing* 26(6): 575–582.

Raškinis, A.; Raškinis G.; Kazlauskienė A. 2004. Lietuvių bendrinės šnekamosios kalbos garsų fondo kūrimo principai, *Lituanistica* 4 (60): 53–62.

Raškinis, G.; Kazlauskienė, A. 2009. Phone frequency in standard Lithuanian. *Respectus philologicus* 16 (21): 169–182.

Sakoe, H.; Chiba, S. 1978. Dynamic programming algorithm optimization for spoken word recognition, *IEEE Transactions on Acoustics, Speech and Signal Processing* 26(1): 43–49.

Soryani, M.; Rafat, N. 2005. Application of Genetic Algorithms to Feature Subset Selection in a Farsi OCR, *Proceedings of the 2005 IEEE Congress on Evolutionary Computation CEC* 3: 2514–2521.

Sotoca, J. M.; Sanchez, J. S.; Mollineda, R. A. 2005. A review of data complexity measures and their applicability to pattern classification problems, *Actas del III Taller Nacional de Mineria de Datos y Aprendizaje. TAMIDA*, 77–83.

Targamadzė, A.; Petrauskienė, R.; Rubliauskas, D. 2010. Influence of technologies on quality of distance learning, *Electronics and Electrical Engineering, Kaunas: Technologija*, 6(102): 131–134.

Theodoridis, S.; and Koutroumbas, K. 1999. *Pattern Recognition*, by Academic Press. 625 p.

Voitovetsky, I.; Guterman, H.; Cohen, A. 1997. Unsupervised speaker classification using self–organizing maps (SOM), *Neural networks for signal processing VII, Processing of the 1997 IEEE Workshop*, 578–587.

Wang, Z.; Hu, W. D; Yu, W. X. 2003. A quick evidential classification algorithm based on k–nearest neighbor rule, *Proceedings of the Second International Conference on Machine Learning and Cybernetics* 5: 3248–3252.

Webb, A. R. 2002. *Statistical Pattern Recognition*. John Wiley, NJ, 2nd edition. 496 p.

Rasa LILEIKYTĖ

ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMAS

Daktaro disertacijos santrauka
Technologijos mokslai, Informatikos inžinerija (07T)

Rasa LILEIKYTĖ

QUALITY ESTIMATION OF SPEECH RECOGNITION FEATURES
DOCTORAL DISSERTATION

Summary of Doctoral Dissertation
Technological sciences, Informatics engineering (07T)