Metrics Based Quality Estimation of Speech Recognition Features

Rasa LILEIKYTĖ*, Laimutis TELKSNYS
Vilnius University, Institute of Mathematics and Informatics
Goštauto 12, LT-01108, Vilnius, Lithuania
e-mail: rasalileikyte@gmail.com, laimutis.telksnys@mii.vu.lt

Abstract. The performance of an automatic speech recognition system heavily depends on the used feature set. Quality of speech recognition features is estimated by classification error, but then the recognition experiments must be performed, including both front-end and back-end implementations. We propose a method for features quality estimation that does not require recognition experiments and accelerate automatic speech recognition system development. The key component of our method is usage of metrics right after front-end features computation. The experimental results show that our method is suitable for recognition systems with back-end Euclidean space classifiers.

Key words: speech recognition, quality of speech recognition features, classes separability.

1. Introduction

The essential and the primary task of developing the automatic speech recognition (ASR) system is to choose the quality feature set. The quality features provide the highest class separability. Thus, the usage of poor features cause ASR system performance degradation.

In general, the ASR system can be divided into two parts: front-end and back-end. The front-end is the part that extracts feature vectors from speech signal. Further, the feature vectors are passed to the back-end classifier. The top level of ASR system architecture is illustrated in Fig. 1. To extract the desired feature system, the speech signal is divided into the overlapping frames. Then, the particular speech extraction algorithm is applied. As a result, feature vectors are derived. Variety of speech recognition feature sets can be constructed including different kinds of features and numbers. After feature extraction, the feature vectors are passed to the back-end classifier of ASR system. Using feature vectors, the back-end of ASR system selects the most consistent word from the reference database. Plenty of ASR research studies have investigated back-end classifiers (Kasparaitis, 2008; Lipeika *et al.*, 2003; Norkin and Keyzer, 2009; Bartkute, 2009). Contrary, we concentrate on feature quality estimation in the front-end of ASR system.

Quality of speech recognition features has been estimated by classification error rate. But then recognition experiments have to be performed with each explored feature set. Recognition experiments need large amount of computational resources.
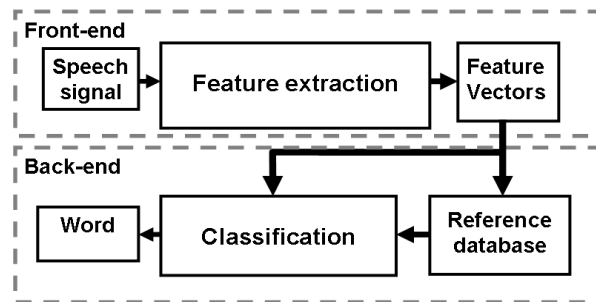
---

*Corresponding author.

Fig. 1. Architecture of the automatic speech recognition system.

The key idea of our paper is to accelerate ASR system development by proposing a method for quality estimation of speech recognition features, that does not require classification error rate computation. Our proposed method is based on metrics employment that describes the level of classes' separability. This approach is suitable for constructing ASR systems with back-end Euclidean space classifiers. It is a group of classifiers that use the Euclidean distance as a similarity measure.

Command based recognition systems are used in various domains, such as home automation, automated vending, warehouse dispatch, order taking and fast serve restaurants, health care monitoring and assisted living, building control, particularly elevators. Moreover, Dynamic Time Warping classifier is Euclidean space classifier and is widely used in command based recognition system due to its simplicity.

In our experiments, we have validated the adequateness of the proposed metrics based method in context of Euclidean space classifiers usage. We have established the quality feature sets for ASR systems with three different back-end Euclidean space classifiers: Dynamic Time Warping (DTW), Nearest Neighbour (NN) and Nearest Centroid (NC). Moreover, our approach has better time complexity of features quality measurement rather than back-end DTW, K–NN classification error rate computation.

This paper is structured as follows. In Section 2, we briefly discuss the related work of features quality estimation. In Section 3, we describe our metrics based method. Section 4 defines the Euclidean space classifiers. The experimental results are shown in Section 5. Finally, Section 6 gives our conclusion.

2. State of the Art

Recently quality of speech recognition features has been estimated by classification error rate. However, this method requires more computational resources. Recognition experiments need large amount of computation resources, including both learning and testing phases implementation. For instance, if ten feature sets are investigated, it means that ten ASR systems have to be created and ten recognition experiments have to be run. It is best visualized by the example in Fig. 2. The quality feature set is chosen with the lowest classification error rate.
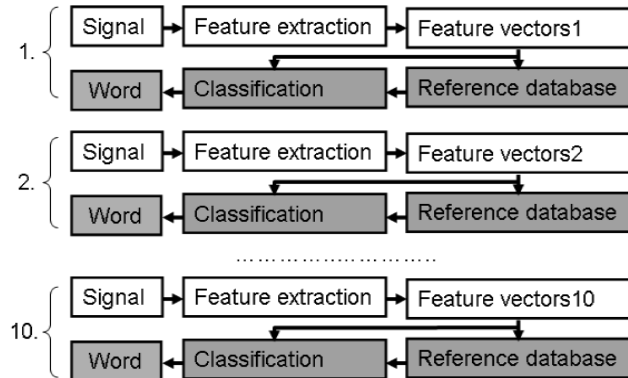
Fig. 2. Traditional method for speech features quality estimation. Shaded regions show classifications with different feature sets.

Table 1
Metrics focused on classes separability.

Group	Metric
Geometrical metrics	Feature efficiency
	Maximal Fisher's discriminant ratio
	Volume of features overlap
	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

Low classification error rate is achieved if classes are well separated. To propose a new method for speech features quality estimation, we also review related metrics concentrated on classes separability and data complexity. The metrics that emphasize on classes separability and data complexity have been recently focused on several aspects: data structure exploration with geometrical metrics (Ho, 2000, 2001, 2002; Ho *et al.*, 2002; Bernado-Mansilla *et al.*, 2005, 2006; Sotoca *et al.*, 2005); synthetic data generation using geometrical metrics (Macia *et al.*, 2008); feature selection, extraction employing information theory metrics (Kerroum *et al.*, 2008; Bao *et al.*, 2006; Chen, 1971) and statistical metrics (Costa *et al.*, 1999; Malina, 1981; Miyamoto *et al.*, 2003). These metrics are presented in Table 1.

Our goal is to propose a method for features quality estimation that does not require performing recognition experiments. We choose back-end Euclidean space classifiers for our research. It is a group of classifiers that use the Euclidean distance. Consequently, the most relevant group of metrics for our task is geometrical metrics as these metrics operate with Euclidean distance. In the studies of other authors, the only study (Mollineda *et al.*, 2005) included experiment for speech signals, but these results were not accented among results of other data kinds. This paper is the continuation of our previous researches. In Lileikytė *et al.* (2011) our objective was to investigate the correlation between geometrical metrics and DTW distance. It was shown that two metrics has the highest correlation with DTW distance: nearest neighbour distances ratio of classes, overstep volume of class boundary. In work Lileikytė *et al.* (2012) we proposed the method for features quality estimation in context of NN classifier usage. The method was based on three metrics usage: features volume of class boundary, nearest neighbour distances ratio of classes and overstep volume of class boundary.

In this paper we: (1) propose the method for features quality estimation that does not require recognition experiments and is based on two metrics computation, (2) confirm experimentally, that our method is suitable for constructing ASR systems with back-end Euclidean space classifiers, including DTW, K-NN, NC classifiers, (3) show that our method accelerates the development of ASR system with back-end DTW, K-NN classifiers. We enable to choose quality speech recognition features in better time complexity.

3. Metrics Based Method for Features Quality Estimation

In this section, we present details of our metrics based method to estimate the quality of speech recognition features.

The quality is the measure that describes the adequacy level of the object's characteristics with determined requirements (ISO 9000, 2000). The high object's quality is achieved if object's characteristics meet the requirements. In context of features quality estimation, the quality feature set gains the lowest classification error rate among other feature sets.

Our metrics based method is focused on constructing ASR systems with particular group of back-end Euclidean space classifiers. It includes the classifiers that use Euclidean distance as a similarity measure.

The classification error rate strongly depends on separability of classes. Classification difficulties encounter on the view that classes are overlapping. The more classes are separable, the lower classification error rate will be. Thus, we concentrate on geometrical characteristics that describe separability of classes with criterion based on the Euclidean distance. For this purpose, we employ two geometrical metrics that gave the best correlation results with DTW distance in our previous study (Lileikytė *et al.*, 2011): nearest neighbour distances ratio of classes, overstep volume of class boundary. In our approach, the classes are well separable if metrics gain minimal values, while classes are overlapped in case of maximal values. Consequently, the quality feature set is identified by low metrics values.
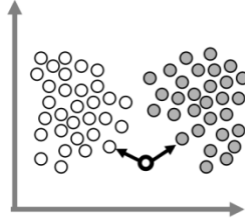
Fig. 3. Distance to the nearest vector of the same and the opposite class.

3.1. *Metrics for Features Quality Estimation*

Nearest neighbour distances ratio of classes. First, for each vector the Euclidean distance is calculated both to the nearest vector of the same class and to the nearest vector of the opposite class (Ho, 2001) (Fig. 3). Then, the ratio is calculated of these distances.

Let $X(1), X(2), \ldots, X(C)$ be the set of classes, where C is the number of classes. Let us define $X(i) = \{x_1(i), x_2(i), \ldots, x_{N_i}(i)\}$ the instance of the i-th class, N_i is vectors number of the i-th class. Let us assume $x_n(i) = \{\boldsymbol{x}_{n1}(i)\boldsymbol{x}_{n2}(i), \ldots, \boldsymbol{x}_{nD}\}$ is the n-th vector of the i-th class, D is the dimension of vector. The metric is computed by the following formula:

$$G1(i, j) = \frac{\sum_{i=1}^{C} \sum_{n=1}^{N_i} \min_k d(x_n(i), x_k(i))}{\sum_{i=1, \, j=1}^{C} \sum_{n=1}^{N_i} \min_m d(x_n(i), x_m(j))}, \tag{1}$$

where $\min_k d(x_n(i), x_k(i))$ is minimal Euclidean distance between the $x_n(i)$ n-th vector of the i-th class and the $x_k(i)$ k-th vector of the i-th class, $1 \leqslant n \leqslant N_i$, $1 \leqslant k \leqslant N_i$, $1 \leqslant i \leqslant C$; $\min_m d(x_n(i), x_m(j))$ is minimal Euclidean distance between the $x_n(i)$ n-th vector of the i-th class and the $x_m(j)$ m-th vector of the j-th class, $1 \leqslant j \leqslant C$, $1 \leqslant m \leqslant N_j$, N_j is vectors number of the j-th class.

Generally, the metric's output is a value between zero and one. This value indicates how much classes separable are, with values closer to zero representing the highest classes separability. Specifically: if the sum of the opposite classes distances are zero, then metric reaches the infinity; if the distance within the same class is bigger than between the opposite classes, the metric's value is bigger than one and not defined. The time complexity of metric algorithm is $O(N \log N)$, where N is the vectors number of all classes.

Overstep volume of class boundary. Assume that every class is represented by the circle. 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 (Li *et al.*, 1988). The radius is computed:

$$r(i) = \max_n d\big(\mu(i), x_n(i)\big), \quad 1 \leqslant n \leqslant N_i, \tag{2}$$

where $\max_n d(\mu(i), x_n(i))$ is maximal Euclidean distance between the $\mu(i)$ mean of the i-th class and the $x_n(i)$ n-th vector of the i-th class, N_i is vectors number of the i-th class $1 \leqslant i \leqslant C$, C is the number of classes.
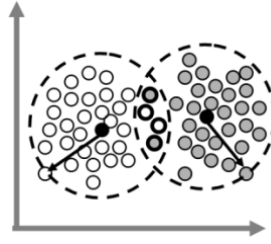
Fig. 4. Vectors of classes falling into the range of circle of the opposite class. Vectors that fall into the range of the opposite class are indicated by heavy borders.

The number of vectors is found that occur in the range of the opposite class (Fig. 4). Then this number is normalized by the vectors number of all classes:

$$G2(i, j) = F\big[d(\mu(i), x_m(j)) \leqslant r(i)\big]\frac{1}{N}, \quad 1 \leqslant m \leqslant N_j, \tag{3}$$

where $d(\mu(i), x_m(j))$ is the Euclidean distance between the $\mu(i)$ mean of the i-th class and the $x_m(j)$ m-th vector of the j-th class, N_j is vectors number of the j-th class, N is the vectors number of all classes. Function $F[\cdot]$ calculates the number of vectors that occur in the area of the opposite class.

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

3.2. *Method for Features Quality Estimation*

We describe and present a proposed metrics based method in Table 2. It includes the steps for metrics calculation and selection of quality feature set. Suppose that we have $X^1, X^2, \ldots, X^P$ feature sets, where P is the number of feature sets, $1 \leqslant p \leqslant P$. Our task is to establish the feature set with the highest quality.

We start by computing the *metrics* $G1^p(i, j)$, $G2^p(i, j)$, as in (**??**) and (**??**) for the each p-th explored feature set. Metrics are calculated for the i-th and the j-th class combination, $1 \leqslant i \leqslant C$, $1 \leqslant j \leqslant C$, C is the number of classes. Then, we compute the mean of metrics results over all classes combinations to reach the *quality indexes* $MG1^p$ and $MG2^p$ that describe the level of quality of the p-th feature set. Quality indexes indicate how classes separable are, with low value representing more separable classes. Finally, we find the *quality feature set* by picking feature set with the lowest quality indexes $MG1$ and $MG2$. The outcome of the metrics based method is the established quality feature set.

Our proposed method for features quality estimation is based on two metrics usage. The method consists of metrics calculation, finding metrics mean and selecting the minimal. Thus, the metrics algorithm needs most computations. Algorithm complexity of nearest neighbour distances ratio of classes is $O(N \log N)$ and of metric overstep volume of class boundary is $O(N)$, where N is vectors number of both reference and test patterns. As

Table 2
Metrics based method to estimate quality of features.

Input: feature sets $\{X^1, X^2, \ldots, X^P\}$, where P is the number of feature sets

1. Compute metric $G1$ as in (1) and metric $G2$ as in (3) for each p-th feature set and each two-class combination (i, j), $1 \leqslant i \leqslant C$, $1 \leqslant j \leqslant C$, $i \neq j$, C is the number of classes

2. Find quality index for each feature set by computing metrics mean

$MG1^p = \sum_{i=1}^{C} \sum_{j=1}^{C} G1^p(i, j) \frac{1}{C^2}, \quad 1 \leqslant p \leqslant P$

$MG2^p = \sum_{i=1}^{C} \sum_{j=1}^{C} G2^p(i, j) \frac{1}{C^2}, \quad 1 \leqslant p \leqslant P$

3. Find the quality feature set: pick feature set having the lowest quality indexes

$MG1^* = \arg\min_p MG1^p, \quad 1 \leqslant p \leqslant P$

$MG2^* = \arg\min_p MG2^p, \quad 1 \leqslant p \leqslant P$

Output: Quality feature set

---

a result, the complexity of the proposed method algorithm is $O(N \log N)$. Also making a notation $N = 2R$, where R is vectors number of speech pattern ($N = R + R$ is a sum of vectors of both reference and test patterns), time complexity of our method algorithm is $O(2R \log 2R)$.

4. Description of Euclidean Space Classifiers

In this section we describe Euclidean space classifiers used in our study. That is classifiers that employ the Euclidean distance as a similarity measure. This group of classifier was chosen because that our metrics based method evaluates the separability of classes in the Euclidean space.

4.1. *Dynamic Time Warping Classifier*

DTW algorithm uses dynamic programming and finds an optimal warping path between two sequences (Sakoe *et al.*, 1978). Suppose we have reference pattern $A = \{a_1, a_2, \ldots, a_R\}$, and test pattern $T = \{t_1, t_2, \ldots, t_Z\}$. The matrix $R \times Z$ is constructed where the (i, j) point is the Euclidean distance $d(a_i, t_j)$ between two vectors a_i and t_j. The l-th point of the warping path W is defined as $w_l = (i_l, j_l)$, where $W = \{w_1, w_2, \ldots, w_L\}$, $\max(R, Z) \leqslant L < R + Z - 1$. We find the path that minimizes the warping cost:

$$D_{\text{DTW}}(A, T) = \min \frac{1}{L} \sum_{l=1}^{L} w_l. \tag{4}$$

This path can be found using dynamic programming algorithm and the optimization criterion:

$$D(i, j) = d(i, j) + \min \left\{ \begin{array}{c} D(i-1, j-1) \\ D(i-1, j) \\ D(i, j-1) \end{array} \right\}, \tag{5}$$

where $D(i, j)$ is the cumulative distance of the optimal path. DTW algorithm is widely used for isolated words recognition. Time complexity of DTW algorithm is $O(R^2)$.

4.2. *K-Nearest Neighbour Classifier*

K-NN is an instance based, supervised learning algorithm (Cover *et al.*, 1967). Suppose we have reference pattern $A = \{a_1, a_2, \ldots, a_R\}$ with labels $Y = \{y_1, y_2, \ldots, y_R\}$, $y_r \in \{\omega_1, \omega_2, \ldots, \omega_C\}$, where $1 \leqslant r \leqslant R$, R is the number of reference vectors, C – number of classes. Let denote test pattern $T = \{t_1, t_2, \ldots, t_Z\}$, where $1 \leqslant z \leqslant Z$, Z is the number of testing vectors. The Euclidean distance is calculated between reference and testing vectors. Testing vector is assigned to the class label with the majority nearest neighbour vectors. In case $k = 1$, we have Nearest Neighbour (NN) classifier. Time complexity of KNN classifier algorithm is $O(K2R \log 2R)$ and of NN classifier it is $O(2R \log 2R)$.

4.3. *Nearest Centroid Classifier*

NC classifier assigns the class label of reference vectors that Euclidean mean is closest to the observed testing vector (Bezdek *et al.*, 2001). Assume $V = \{v_1, v_2, \ldots, v_C\}$ denotes a set of centroids, where C is the number of classes. The centroid is computed as the Euclidean mean of class vectors. Let $T = \{t_1, t_2, \ldots, t_Z\}$ be a test pattern. The NC classifier assigns the testing vector t_z to the class of its nearest centroid v_i, $1 \leqslant i \leqslant C$. Time complexity of NC classifier algorithm is $O(2R \log 2R)$.

5. Results and Discussions

In this section we present experimental results of our proposed metrics based method. Also we provide validation results of this method. The experiments are constructed on Lithuanian speech corpus VDU-TRI4 (Raškinis *et al.*, 2004). The fourteen most frequent Lithuanian phonemes are selected for this experimental study (Raškinis *et al.*, 2009): /a/, /e/, /i/, /j/, /k/, /m/, /n'/, /o:/, /r/, /r'/, /s/, /s'/, /t/, /t'/ (n', t', s' are soft consonants and o: is long vowel). The phonemes are extracted from the speech recordings of four speakers, including two males and two females. The sampling rate of speech corpus is 11025 Hz, 16-bits/sample, mono-channel. Each class consists of 300 instances. The experiments are made for 91 two-class combinations.

We build the metrics based system as it is described in Section 3. Two feature sets are under investigation: 12th order Linear Frequency Cepstral Coefficients (LFCC), 12th
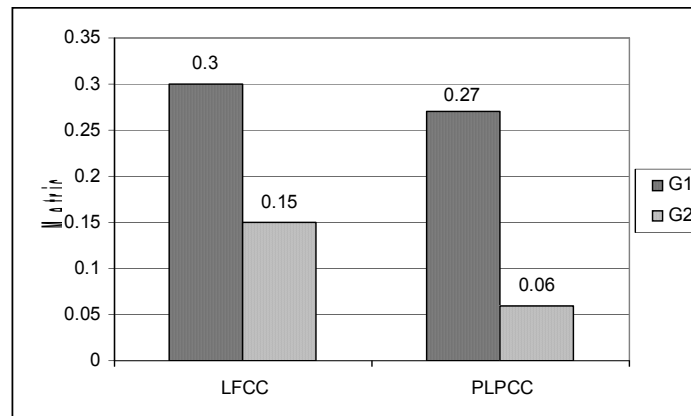
Fig. 5. Quality indexes of two different feature sets.

order Perceptual Linear Prediction Cepstral Coefficients (PLPCC). Figure 5 shows the evaluation results of quality indexes *G1* and *G2* for this experiment.

Our results indicate that PLPCC feature set outperforms the feature set of LFCC. As it can be seen from Table 3, the lowest quality indexes of *G1* and *G2* are achieved for PLPCC features rather than LFCC. In case of quality index *G1*, the small improvement is observed of three units (0.30–0.27). A slight difference shows that the degree of classes' separability in the view of distance between vectors of the same class and opposite class is similar for both feature sets. The quality index *G2* results improvement is gained of nine units (0.15–0.06). It can be seen, that index *G2* of PLPCC features is more than twice bigger than LFCC. This suggests that classes of PLPCC feature set are situated more compact and have smaller radius than LFCC.

To validate the adequateness of the proposed metrics based method, we report experimental results of classification error rate on different kinds of ASR systems. Six kinds of ASR systems for each data set are constructed, consisting of two different front-end feature sets: (1) 12th order PLPCC, (2) 12th order LFCC and three different back-end classifiers: (1) DTW, (2) NN, (3) NC. The classifiers are described in Section 4. As well as this, we evaluate the classification error rate of six different ASR systems in case of three explored data sets and the results are given in Table 3.

From Table 3 one can infer that the ASR systems with front-end PLPCC features performed best compared to LFCC features.

Conclusions can be drawn from Fig. 5 and Table 3 that the experimental results validated the adequateness of the proposed metrics based method in context of Euclidean space classifiers usage. Employing both the metrics based method and the classification error rate, it is established that PLPCC feature set is the quality feature set. This research demonstrates that quality of speech features can be measured right after ASR front-end execution, applying metrics based method. It was showed that metrics based method is a good indicator of evaluating the efficiency of features with lesser resources comparing to classification error rate computation. Our metrics based method achieves better time

Table 3
Recognition error rates (%) of six different recognition systems.

Recognition system		Error rate (%)
Feature set	Classifier	
LFCC	DTW	5.51 ± 0.86
PLPCC	DTW	$\mathbf{3.21 \pm 0.67}$
LFCC	NN	13.42 ± 1.52
PLPCC	NN	$\mathbf{9.12 \pm 1.07}$
LFCC	NC	23.81 ± 1.65
PLPCC	NC	$\mathbf{18.55 \pm 1.25}$

complexity over the classification error rate based method for DTW and K-NN classifiers. Time complexity of metrics based method is $O(2R \log 2R)$, where R is vectors number of speech pattern. While time complexity of DTW algorithm is $O(R^2)$ and K-NN algorithm requires $O(K \, 2R \log 2R)$.

6. Conclusion

A new metrics based method for quality estimation of speech recognition features was proposed in this paper. The previous traditional method for speech features quality estimation is based on recognition error rate computation and requires performing recognition experiments, including both front-end feature extraction and back-end classifier implementations. In contrast to it, our method measures the quality right after front-end feature extraction and does not require performing back-end classifier experiments. Our method includes two metrics computation: nearest neighbour distances ratio of classes, and overstep volume of class boundary. The metrics define the level of classes separability in the Euclidean space.

The experimental results demonstrated that the proposed method is suitable for features quality estimation when constructing automatic speech recognition systems with back-end Euclidean space classifiers, including Dynamic Time Warping, Nearest Neighbour, and Nearest Centroid classifiers. Our method improves time complexity of speech features quality estimation in context of Dynamic Time Warping, K-Nearest Neighbors classifier usage.

References

Bao, M., Guan, L., Li, X., Tian, J., Yang, J. (2006). Feature extraction using histogram entropies of Euclidean distances for vehicle classification. In: *International Conference on Computational Intelligence and Security*, vol. 1, pp. 668–673.

Bartkutė-Norkienė, V. (2009). Stochastic optimization algorithms for support vector machines classification. *Informatica*, 20(2), 173–186.

Bernado-Mansilla, E., Ho, T.K., Orriols, A. (2006). Data complexity and evolutionary learning. In: *Data Complexity in Pattern Recognition*. Springer, Berlin, pp. 115–134.

Bezdek, J.C., Kuncheva, L.I. (2001). Nearest neighbour classifier designs: an experimental study. *International Journal of Intelligent Systems*, 16, 1445–1473.

Chen, C.H. (1971). Theoretical comparison of a class of feature selection criteria. *IEEE Transaction on Computers C, Pattern Recognition* 209, 1054–1056.

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* 4711, 3108–3115.

Cover, T.M., Hart, P.E. (1967). Nearest neighbour pattern classification. *IEEE Transactions on Information Theory IT*, 131, 21–27.

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

Ho, T.K. (2001). Data complexity analysis for classifier combination. In: *Multiple Classifier Systems*, vol. 2096. Springer, Berlin, pp. 53–67.

Ho, T. K. (2002). A data complexity analysis of comparative advantages of decision forest constructors. *Pattern Analysis and Applications,* 5, pp. 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.

ISO 9000 (2000). *Quality Management Systems – Fundamentals and Vocabulary*.

Kasparaitis, P. (2008). Lithuanian speech recognition using the English recognizer. *Informatica*, 19(4) 505–516.

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

Li, Z., Ditang, F. (1988). The research on speech feature representation method and distance measure method. In: *9th International Conference on Pattern Recognition*, vol. 1, pp. 631–633.

Lipeika, A., Lipeikienė, J. (2003). Word endpoint detection using dynamic programming. *Informatica*, 14(4), 487–496.

Lileikytė, R., Telksnys, L. (2011). Quality estimation methodology of speech recognition features. *Electronics and Electrical Engineering*, 4(110), 113–116.

Lileikytė, R., Telksnys, L. (2012a). Quality measurement of speech recognition features in context of nearest neighbour classifier. *Electronics and Electrical Engineering*, 2(118), 9–12.

Lileikytė, R., Telksnys, L. (2012b). Quality estimation of speech recognition features for dynamic time warping classifier. *Information Technology and Control*, 41(110), 268–273.

Macia, N., Bernado-Mansilla, E., Orriols-Puig, A. (2008). Preliminary approach on synthetic data sets generation based on class separability measure. In: *IEEE International Conference on Pattern Recognition*, pp. 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.

Mollineda, R.A., Sanchez, J.S., Sotoca, J.M. (2005). Data characterization for effective prototype selection. In: *First Edition of the Iberian Conference on Pattern Recognition and Image Analysis (IbPRIA 2005). Lecture Notes in Computer Science*, vol. 3523. Springer, Berlin, pp. 27–34.

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

Norkin, V., Keyzer, M. (2009). On stochastic optimization and statistical learning in reproducing kernel Hilbert spaces by support vector machines (SVM). *Informatica*, 20(2), 273–292.

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

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

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.

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

R. Lileikytė doctor in informatics of Vilnius University, Institute of Mathematics and Informatics. Her research interests are in speech processing, including speech recognition, signal processing, and speech synthesis.

L. Telksnys professor, doctor habilitatis in informatics, doctor honoris causa of Kaunas University of Technology, member of Lithuanian Academy of Sciences, researcher at Recognition Processes Department at Vilnius University, Institute of Mathematics and Informatics, Vilnius, Lithuania. He is the author of an original theory of detecting changes in random processes, investigator and developer of a computerized system for statistical analysis and recognition of random signals. His current research interests are in analysis and recognition of random processes, cardiovascular signals and speech processing and computer networking.

Metrikomis grindžiamas šnekos atpažinimo požymių kokybės vertinimas

Rasa LILEIKYTĖ, Laimutis TELKSNYS

Automatinės šnekos atpažinimo sistemos tikslumas priklauso nuo naudojamų požymių. Šnekos atpažinimo požymių kokybė yra vertinama paskaičiuojant klasifikavimo klaidą. Tačiau tokiu atveju turi būti atliekami atpažinimo eksperimentai, vykdant požymių išskyrimo bei klasifikavimo modulius. Mes sukūrėme metodą požymių kokybei vertinti, kuris nereikalauja atpažinimo eksperimentų vykdymo ir paspartina atpažinimo sistemos kūrimą. Svarbiausias sukurto metodo komponentas yra metrikų naudojimas po požymių išskyrimo modulio. Eksperimentiniai rezultatai parodė, kad šis metodas yra tinkamas konstruoti atpažinimo sistemas su Euklido erdvės klasifikatoriais.