DIRECT EVALUATION OF FRAME-BASED NONSTATIONARY PATTERN RECOGNITION METHODS BY USING BHATTACHARYYA DISTANCE

Milan Marković[1], *Milan Milosavljević*[1,2], *Branko Kovačević*[2]
[1]Institute of Applied Mathematics and Electronics
Kneza Miloša 37, 11000 Belgrade, Yugoslavia
e-mail: emarkovm@ubbg.etf.bg.ac.yu
[2]Faculty of Electrical Engineering, University of Belgrade
Bulevar Revolucije 73, 11000 Belgrade, Yugoslavia
e-mail: emilosam@ubbg.etf.bg.ac.yu, Fax: +381 11 324 86 81

ABSTRACT

In this paper, a possibility of evaluating frame-based nonstationary pattern recognition methods by using Bhattacharrya distance is considered. Speech signal is used as a nonstationary signal and the comparative analysis is done through analyzing the natural speech, isolately spoken serbian vowels and digits.

1 INTRODUCTION

Bayes error is a very important performance index in pattern recognition, assessing the classifiability of data and measuring the discrimination capabilities of the features, even before considering what type of classifier should be designed. However, the calculation of the error probability is a very difficult task. Even when observation vectors have a normal distribution, we must resort to numerical techniques [1]. However, a closed-form expression for the Bayes error probability is the most desirable solution for a number of reasons. When we cannot obtain a closed-form expression, we may seek either an Bayes error estimate, or an upper bound of the Bayes error probability. As for the Bayes error estimation, very efficient estimation procedures based on *k*-NN approach are proposed in [1,2,3].

A situation is more complicated in case of the nonstationary pattern recognition methods. Namely, statistical pattern recognition methods are based on the assumption of stationarity of the processes to be recognized. There are many problems in applying these methods in the real-time recognition of data obtained from the nonstationary processes. The main problems are limited validity and size of the learning data set. One approach for solving these problems represents the frame-based pattern recognition methods. These methods are based on the idea that the signal from the nonstationary data process should be considered in frames [4], and on using the unsupervised learning procedures for classifier design on the given frame of signal and its application as the initial classifier for the next frame. Based on the obtained initial partition of the next frame, the same unsupervised learning procedure gets started, and so on. As the unsupervised learning algorithm, the c-mean clustering algorithm or nearest mean reclassification rule is proposed in [4]. After that, the iterative application of quadratic classifier, described by Fukunaga [1], as a more sophisticated and efficient procedure is proposed, and its application as an unsupervised learning procedure in the frame-based non-stationary pattern recognition systems is justified in [5]. Additionally, in [6], a modification of the proposed iterative quadratic classifications procedure with the increased real-time application possibilities is proposed. It is shown in [6], that proposed modification has some robust characteristics to the inappropriateness of the assumed classification model.

This paper is dedicated to possible finding of an optimal criterion for evaluation the considered frame-based nonstationary pattern recognition methods: the c-mean clustering algorithm with Euclidean distance, the iterative quadratic classifications, and its real-time modification. Due to the fact that these methods are based on the unsupervised principle, we believe that the complicated *k*-NN Bayes error estimation procedure, considered in [1,2,3], is not suitable for the nonstationary cases. Instead of this, in order to evaluate the considered frame-based nonstationary pattern recognition methods we propose the use of the upper bound trajectories of the Bayes error obtained by using Bhattacharyya distance. The main purpose of this analysis lies in the investigating the possibility of using upper bound trajectories of the Bayes error obtained by using Bhattacharyya distance in direct evaluation of the frame-based nonstationary pattern recognition method. In the paper, the speech is used as a nonstationary signal and the comparative analysis is done through analyzing the natural speech, isolately spoken serbian vowels and digits. In fact, the considered methods are applied in a combined nonrobust/robust recursive AR speech analysis procedure, proposed in [6].

Paper is organized as follows. Section 2 is dedicated to brief description of Bhattacharyya distance. Experimental analysis is presented in Section 3 while conclusion is given in Section 4.

2 BHATTACHARYYA DISTANCE

Bayesian classifier could be described by follows. Let us consider c classes of the training data set, ω_i, $i=1,\dots,c;$

described by a posteriori probability functions $P(\omega_i|X)$. Bayes rule could be expressed as follows, [1]:

$$P(\omega_i|X)=\frac{P(\omega_i)p(X|\omega_i)}{p(X)} \tag{1}$$

where: $p(X)$ is a probability density function (pdf) of sample X, $p(X|\omega_i)$ is conditionally probability density function (cpdf), and $P(\omega_i)$ is a priori probability of the class ω_i. Bayesian classifiers is based on *Bayes decision rule* which is referred to the classification of sample X in that class ω satisfying:

$$P(\omega|X)=\max_{1\le i\le c}\{P(\omega_i|X)\} \tag{2}$$

The Bayes error is given by:

$$\begin{aligned} e &= 1-\int_{R^n}\max_{1\le i\le c}P(\omega_i|X)p(X)dx \\ &= 1-\int_{R^n}\max_{1\le i\le c}P(\omega_i)p(X|\omega_i)dx \end{aligned} \tag{3}$$

In a two-class case (c=2), the equation (3) could be rewritten as:

$$e=\int\min\left[P_1p_1(X),P_2p_2(X)\right]dX \tag{4}$$

where: P_1, P_2, and $p_1(X)$, $p_2(X)$ denote a posteriori probabilities and cpdfs of the first and second class, respectively. An upper bound of the integrand in (4) may be obtained by making use of the fact that:

$$\min[a,b]\le a^s b^{1-s}\,,\, 0\le s\le 1\,,\; for\; a,b\ge 0 \tag{5}$$

Using the inequality (5), e can be bounded by:

$$e_u=P_1^sP_2^{1-s}\int p_1^s(x)p_2^{1-s}(X)dX \quad for \quad 0\le s\le 1 \tag{6}$$

where e_u indicates an upper bound of e. This e_u is called the Chernoff bound [7]. The optimum s can be found by minimizing e_u. When two density functions are normal, $N_X(M_1,\Sigma_1)$ and $N_X(M_2,\Sigma_2)$, the integration (6) can be carried out to obtain a closed-form expression for e_u. That is:

$$\int p_1^s(X)p_2^{1-s}(X)dX=e^{-\mu(s)} \tag{7}$$

$$\begin{aligned}\mu(s)=&\frac{s(1-s)}{2}(M_2-M_1)^T\left[s\Sigma_1+(1-s)\Sigma_2\right]^{-1}(M_2-M_1)+\\ &+\frac{1}{2}\ln\frac{|s\Sigma_1+(1-s)\Sigma_2|}{|\Sigma_1|^s|\Sigma_2|^{1-s}}\end{aligned} \tag{8}$$

The expression (8) for $\mu(s)$ is called the Chernoff distance. For this case, the optimum s can be easily obtained by plotting $\mu(s)$ for various s with given M_i and Σ_i. The optimum s is the one which gives maximum value for $\mu(s)$.

In case that we do not insist on the optimum selection of s, we may obtain a less complicated upper bound. One of the possibilities is to select s=1/2. Then the upper bound is:

$$e_u=\sqrt{P_1P_2}\int\sqrt{p_1(X)p_2(X)}\,dX=\sqrt{P_1P_2}\;e^{-\mu(1/2)} \tag{9}$$

in general, and for normal distributions:

$$\begin{aligned}\mu(1/2)=&\frac{1}{8}(M_2-M_1)^T\left[\frac{\Sigma_1+\Sigma_2}{2}\right]^{-1}(M_2-M_1)+\\ &+\frac{1}{2}\ln\frac{\left|\frac{\Sigma_1+\Sigma_2}{2}\right|}{\sqrt{|\Sigma_1||\Sigma_2|}}\end{aligned} \tag{10}$$

The term $\mu(1/2)$ is called the Bhattacharyya distance, and will be used as an important measure of the separability of two distributions [8]. In fact, the Bhattacharyya distance is the optimum Chernoff distance when $\Sigma_1=\Sigma_2$. As seen in (10), the Bhattacharyya distance consists of two terms. The first or second terms disappears when $M_1=M_2$ or $\Sigma_1=\Sigma_2$, respectively. Therefore, the first term gives the class separability due to the mean-difference, while the second term gives the class separability due to the covariance difference. It is important to know which term is dominant, because it determines what type of a classifier must be designed for given distributions.

3 EXPERIMENTAL ANALYSIS

The test signal consists of five isolately spoken vowels (“a”, “e”, “i”,“o”,”u”) and ten isolately spoken digits (“1”, “2”, ... ,”0”) from one speaker. The signal is sampled with f_S=10kHz and preemphasized with q=1. All experimental results are obtained by using AR model of *10*th order. As the objective quality measure, the MAR (Mean Absolute Residual) criterion is used [6]:

$$J=1/\mathrm{M}\cdot\sum_{i=1}^{\mathrm{M}}|\mathrm{s(i)}-\hat{\mathrm{s}}(\mathrm{i})| \tag{11}$$

where $s(i)$ is the speech sample at the i-th time instance, $\hat{s}(i)$ is its linear prediction, and M is total number of processed speech samples.

In this paper, we consider the nonrobust/robust recursive estimation procedure for parameter identification of nonstationary AR speech model based on the weighted recursive least squares algorithm (WRLS) with variable forgetting factor (VFF) and unsupervised learning procedures for design of frame-based nonstationary pattern classifier. We consider the following unsupervised learning methods: the c-mean clustering algorithm with Euclidean distance (denoted as CEUC), the iterative quadratic classifications (CIQC), and its real-time modification (RTQC). As a direct comparative criteria of the considered methods (CEUC, CIQC, and RTQC algorithms), the upper bound trajectories of the Bayes error obtained by using Bhattacharyya distance [8] (which values are calculated at each time instances on the basis of the estimated corresponding classifier parameters) are used. The results of this evaluation is compared to the results obtained by indirect evaluation of the considered methods [5,6].

As the examples, the estimated upper bound trajectories of the Bayes error based on Bhattacharrya distance, obtained by using the CEUC, CIQC, and RTQC algorithms in analyzing the vowel “A” and digit “1” are presented on Fig. 1 and 2. As the indirect comparative criteria, we use adaptiveness, sensitivity to the pitch impulses, bias, and variance of the obtained trajectories of the non-stationary AR speech model parameter estimates. Fig. 3 and 4 show the estimated trajectories of the first AR parameter (AR_1) obtained by using the proposed robust recursive procedure with application of considered frame-based procedures (frame length was 100 samples): CEUC, CIQC, and RTQC algorithms in analyzing the vowel “A” and digit “1”,

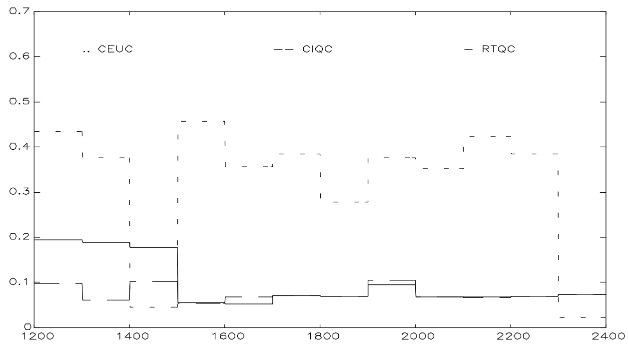

Fig.1: Upper bound trajectories of Bayes error obtained by using: CEUC, CIQC, and RTQC algorithms in analyzing the vowel “A” (frame length is N=100)

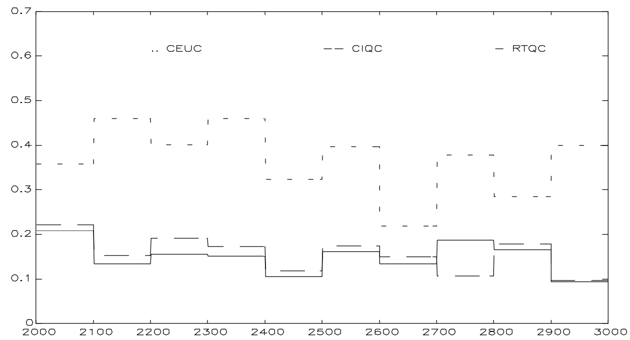

Fig.2: Upper bound trajectories of Bayes error obtained by using: CEUC, CIQC, and RTQC algorithms in analyzing the digit “1” (frame length is N=100)

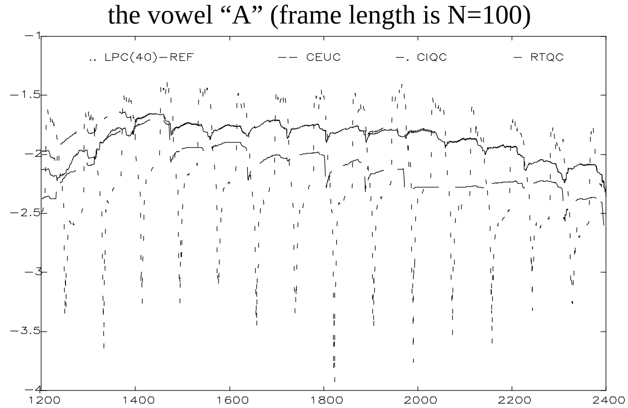

Fig.3: The AR_1 parameter trajectories obtained by using: LPC(40)-REF, CEUC, CIQC, and RTQC algorithms in analyzing the vowel “A” (frame length is N=100).

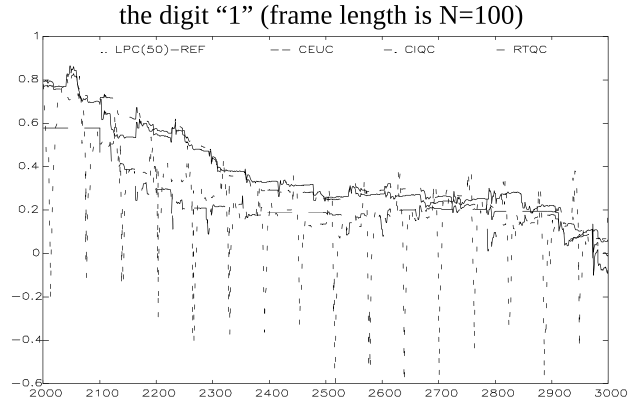

Fig.4: The AR_1 parameter trajectories obtained by using: LPC(50)-REF, CEUC, CIQC, and RTQC algorithms in analyzing the digit “1” (frame length is N=100).

respectively, for the signal frames showed on Fig. 1 and 2. In this case, we use reference parameter trajectories and comparative methodology, proposed in [9].

As for the experimental results obtained through analyzing of the all test data files, the mean values of the upper bound Bayes error trajectories obtained by using the proposed methods in analysis of vowels and digits are presented in Table 1. Also, to evaluate which term of the Bhattacharyya distance (10) is dominant, examples of trajectories of the distance, first and second term of (10) are given on Fig. 5.

Based on experimental results, presented in Table 1 and Fig. 1, and 2, we could conclude that lower upper bounds of the Bayes error probability, e.g. better results in both vowels and digits analysis, are obtained by using the proposed robust recursive procedure with application of the unsupervised learning procedure for classifier design based on quadratic classifier (CIQC and RTQC algorithm). This corresponds well to the comparative results, presented on Fig. 3, and 4, showing that the trajectories of AR_1 parameter estimates obtained by the robust recursive AR speech procedure with application of CIQC, and RTQC algorithms have lower bias, lower variance, more adaptiveness to the nonstationarity of the model parameters, and lower sensitivity to the pitch impulses than the same robust recursive procedure with application of CEUC algorithm for classifier design.

Table 1: Mean values of the upper bounds of Bayes error; vowels and digits analysis.

F	Length	CEUC	CIQC	RTQC
A	3690	0.336	0.160	0.164
E	3690	0.335	0.150	0.143
I	3690	0.340	0.141	0.150
O	3690	0.332	0.140	0.191
U	3690	0.322	0.157	0.180
1	6690	0.325	0.194	0.230
2	6690	0.331	0.202	0.202
3	5690	0.332	0.211	0.232
4	6690	0.326	0.192	0.243
5	6690	0.338	0.202	0.245
6	7690	0.333	0.209	0.237
7	6690	0.321	0.198	0.228
8	7690	0.321	0.208	0.236
9	5690	0.325	0.225	0.218
0	5690	0.293	0.196	0.207

The example, presented on Fig. 5, shows that the second term of the Bhattacharyya distance (10), is dominant meaning that some classifier with nonlinear discrimination function (such as quadratic classifier) should be used instead of classifier with linear discrimination function (such as c-mean algorithm with Euclidean distance).

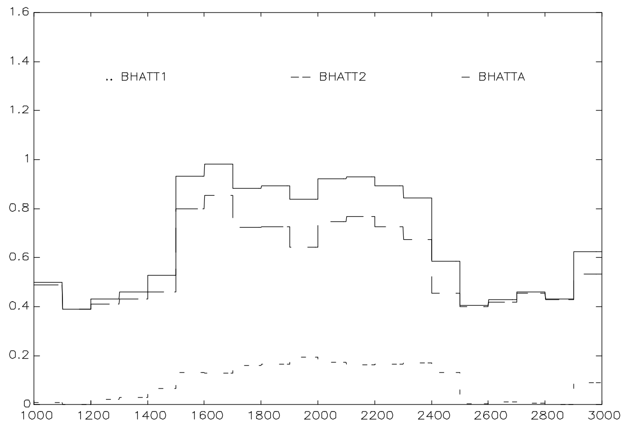

Fig. 5: Bhattacharyya distance values obtained in analyzing the vowel "A"; BHATTA, BHATT1, and BHATT2 represent the Bhattacharyya distance, first term, and second term of the (10), respectively.

As for the direct comparative analysis of the CIQC and RTQC algorithm, the presented results show that both algorithms produce similar results with slightly lower mean values of the upper bounds of the Bayes error obtained by using the CIQC algorithm, presented in Table 1. Based on the entire experimental analysis, it could be concluded that the trajectories of the upper bound of Bayes error based on Bhattacharrya distance could be used as a direct evaluation criterion for evaluating the frame-based nonstationary pattern recognition methods.

4 CONCLUSION

In this paper, we consider a possibility of direct evaluation of frame-based nonstationary pattern recognition methods by using Bhattacharrya distance. The nonstationary pattern recognition methods based on unsupervised learning procedures: c-mean clustering, iterative quadratic classifications, and its modification for real-time purposes, are considered. The considered methods are evaluated through their applications for design of nonstationary pattern classifier in nonrobust/robust AR speech parameter estimation procedure based on the weighted recursive least squares algorithm with variable forgetting factor. The comparative analysis is done through analyzing the natural speech, isolately spoken serbian vowels and digits. Based on experimental results, we could conclude that lower upper bounds of the Bayes error probability, e.g. better results in both vowels and digits analysis, are obtained by using the proposed estimation procedure with application of procedure for classifier design based on quadratic classifier, compared to the c-mean clustering algorithm. This corresponds well to the comparative results based on the estimated parameter trajectories obtained by the considered robust recursive AR speech estimation procedure. Based on the entire experimental analysis, it could be concluded that the trajectories of the upper bound of Bayes error based on Bhattacharrya distance could be used as a direct evaluation criterion for the frame-based nonstationary pattern recognition methods.

ACKNOWLEDGMENTS

This paper was supported by Ministry of Sciences and Technology of Republic Serbia, projects No. 04M02 and No. 1007, through Institute of Mathematics SANU Belgrade and Faculty of Electrical Engineering, University of Belgrade, respectively.

REFERENCES

[1] K.Fukunaga, *Introduction to Statistical Pattern Recognition*, ACADEMIC PRESS INC., 1990.

[2] K.Fukunaga and D.Hummels, "Bayes Error Estimation Using Parzen and *k*-NN Procedures," *IEEE Trans. on PAMI*, vol. PAMI-9, pp. 634-643, 1987.

[3] Lj.Buturović and M.Marković, "Improving *k*-Nearest Neighbor Bayes Error Estimates," in *Proc. of 11th IAPR International Conference on Pattern Recognition*, The Hague, The Netherlands, August 30-September 3, 1992.

[4] D.Gutfinger and J.Sklansky, "Robust Classifiers by Mixed Adaptation," *IEEE Trans. on PAMI*, Vol. 13, No. 6, June 1991., pp. 552-567.

[5] M.Marković, M.Milosavljević, B.Kovačević, "Clustering in Non-Stationary Pattern Recognition Systems," *SIGNAL PROCESSING VII: Theories and Applic., M. Holt, C. Cowan, P. Grant, W. Sandham (Eds.), 1994. European Association for Signal Processing* (*Proc. of EUSIPCO-94*, Edinburgh, Scotland, 13-16 September, 1994.) pp. 1803-1806.

[6] M.Marković, B.Kovačević, M.Milosavljević, "A Statistical Pattern Recognition Approach to Robust Recursive Identification of Non-Stationary AR Model of Speech Production System," *IEEE Trans. on Speech and Audio Processing*, Vol. 4, No. 6, November 1996, pp. 456-460.

[7] H.Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Ann. Math. Stat.*, 23, pp. 493-597, 1952.

[8] A.Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, 35, pp. 99-110, 1943.

[9] M.Veinović, B.Kovačević, M.Milosavljević, "Robust Non-recursive AR Speech Analysis", *Signal Processing*, Vol. 37, No. 2, May 1994., pp. 189-201.