Yugoslav Journal of Operations Research
6 (1996), Number 1, 113-120

ITERATIVE QUADRATIC CLASSIFICATIONS IN NON-STATIONARY PATTERN RECOGNITION SYSTEMS

Milan Ž. MARKOVIĆ [1], Milan M. MILOSAVLJEVIĆ [1]
and
Branko D. KOVAČEVIĆ [2]

[1] *Institute of Applied Mathematics and Electronics, Kneza Miloša 37, 11000 Belgrade, Yugoslavia*
[2] *University of Belgrade, Faculty of Electrical Engineering, Bulevar Revolucije 73, 11000 Belgrade, Yugoslavia, Fax: (381-11) 324-8681*

Abstract. In this paper we propose the use of an iterative quadratic classifications algorithm as unsupervised training procedure in the frame-based approach to the recognition of non-stationary patterns. We present a comparative experimental analysis of the proposed algorithm and the well-known c-mean clustering algorithm. The efficiency of these two iterative clustering procedures is experimentally evaluated through their application in the robust recursive identification of the time-varying autoregressive (AR) model of speech signals. Means and standard deviations of the Mean Absolute Residual (MAR) criterion as well as trajectories of the estimated AR parameters show the superiority of the proposed procedure over the c-mean clustering algorithm.

Keywords: Iterative quadratic classifications, non-stationary pattern recognition systems, robust recursive AR speech analysis.

1. INTRODUCTION

Statistical pattern recognition methods are usually based on the assumption that the processes to be recognized are stationary. There are many problems in applying these methods to the recognition of a large amount of data obtained from non-stationary processes. The main problems are limited validity and the size of the training data set. In other words, due to data that are not stationary, the training data set is valid only for a limited (unknown) number of data samples so that some adaptation of classifier parameters must be made. Also, the choice of the training data set size is problematic. Namely, a large training data set does not guarantee the non-stationary data will be effectively dealt with and a small training data set does not guarantee the unbiased estimation of the classifier parameters. Due to their complexity, these problems are very little elaborated in the literature. There are only a few ideas about the ways to find solutions.

One of these ideas is that the signal from the non-stationary data process should be considered in *frames* [1]. Namely, the proposed method consists of an iterative unsupervised training procedure to design the classifier for the given frame of signal and its application as the initial classifier for the next frame. Based on the initial partition of the next frame, the same iterative unsupervised training procedure gets started, and so on. Gutfinger & Sklansky [1] proposed a c-mean clustering algorithm or nearest mean reclassification rule as the unsupervised training procedure.

In this paper, we propose a more sophisticated clustering procedure to use in the mentioned frame-based non-stationary pattern recognition approach. The proposed method consists of the *iterative application of the quadratic classifier* [2] as the unsupervised training procedure for classifier design on each frame of the signal. Namely, a final quadratic classifier for each frame of the signal is obtained from the initial quadratic classifier for that frame through an iterative procedure of quadratic classifications. The final quadratic classifier for each frame represents the initial classifier for the next frame of the signal, and so on. The efficiency of the proposed method is experimentally evaluated and compared to the c-mean algorithm through their application in the robust recursive identification of the time-varying AR model of speech signals.

The paper is organized in the following way. The proposed method is described in Section 2. Section 3 is dedicated to the application of the proposed method in the robust recursive time-varying AR speech analysis. A comparative experimental analysis is presented in Section 4. Conclusions and a summary are provided in Section 5.

2. DESCRIPTION OF THE ALGORITHM

Just like the c-mean algorithm, the iterative quadratic classification clustering algorithm is derived from the general clustering algorithm, described in [2]. Assume that we want to classify N samples, $X_1, \ldots, X_N$ into one of L classes, $w_1, \ldots, w_L$, where L is assumed to be given. The iterative quadratic classification clustering algorithm has the form:

Step 1. Choose an initial partition of given data set and calculate: $P_i(0)$ (a priori class probability), $M_i(0)$ (mean class vectors), and $\Sigma_i(0)$ (a class covariance matrix) for $i=1, \ldots, L$.

Step 2. Having calculated a priori class probabilities, $P_i(l)$, mean vectors, $M_i(l)$, and covariance matrices, $\Sigma_i(l)$, at the l-th iteration, reclassify each X_j according to the smallest $(1/2)(X_j-M_i)^T \Sigma_i^{-1}(X_j-M_i) + (1/2) \ln |\Sigma_i| \ln P_i$. The a priori class probability for w_i is estimated by the number of w_i-samples divided by the total number of samples.
If the classification of any X_j is changed, calculate $P_i(l+1)$, $M_i(l+1)$, and $\Sigma_i(l+1)$ for the new class assignment, and go to Step (2). Otherwise, stop.

The application of the proposed algorithm in recognizing non-stationary data could be described by the following. The final quadratic classifier for each frame of siggnal is used as the initial quadratic classifier for the next frame to produce its initial partition. Based on that initial partition the same iterative clustering procedure starts to produce the final quadratic classifier for that frame, and so on. The initial partition for the initial frame of the signal is chosen heuristically.

Similar to the consideration in [1], the convergence of the proposed algorithm depends on the following atributes of the initial calssifier and the data: the error rate of the initial classifier, the size of the signal frames, and the validity of the model used to represent the class conditional probability densities.

The efficiency of the two iterative clustering procedures is experimentally evaluated through their application in the robust recursive identification of the time-varying AR model of speech signals.

3. APPLICATION IN RECURSIVE AR SPEECH ANALYSIS

In this work, speech is used as an example of a non-stationary signal. In fact, the two iterative clustering procedures are applied in a combined non-robust/robust recursive AR speech analysis procedure [3] based on RLS algorithm with variable forgetting factor (VFF) to classify the residual speech samples into two classes. The first class consists of "small" residual samples and the second one consists of "big" residual samples. The classification of the k-th residual sample selects either the non-robust or the robust recursive AR procedure for LPC parameter estimation at the $(k+1)$-time instance. This method is based on the well-known assumption regarding excitation for voiced speech as an innovative process from the mixture distribution, such that a large portion of the excitations are from a normal distribution with a very small variance while a small portion of the glottal excitations are from an unknown distribution with a much bigger variance [4]. In this case, the classifier is very simple, one-dimensional, and mean vectors and covariance matrices are means and variances, respectively. The classification consists of two steps: *initialization* and *adaptation*.

Initialization: On the initial segment of the signal, the initial classifier is obtained applying either the c-mean clustering procedure with Euclidean distance (CEUC algorithm) or the iterative quadratic classification procedure (CIQC algorithm) based on an initial partition of the initial segment that is heuristically chosen.

Adaptation: The initial classifier is applied in the classification of the residual speech samples obtained in the proposed recursive AR speech analysis procedure on the next segment of the signal with N sample size. The result of the k-th residual sample classification selects either the non-robust recursive procedure (first class) or robust recursive procedure (second class) to estimate the vector of AR parameters in the k-time instance. The obtained vector is used to determine the $(k+1)$ residual sample and the procedure continues. The classification result of the entire segment represents the initial partition of that segment and is used to start the given iterative clustering procedure to produce the initial classifier for the next segment of N residual samples, and so on.

A heuristic is defined to keep this classifier in a stable state. Namely, if all residual samples of a given segment are classified into the same class (first or second) then some special, "neutral", value of the classifier parameters (heuristically or experimentally defined) is used to define the classifier for the next segment of the signal.

The convergence property of the proposed robust estimation algorithm is mainly determined by the conventional RLS algorithm with VFF (non-robust recursive procedure), since the a priori probability of the first class is in most cases significantly greather than the a priori probability of the second one (in the case of voiced speech the typical values are 0.9 and 0.1 respectively). The robust part of the proposed procedure improves the convergence properties, since the robust RLS procedure suppress the influence of the spiky parts of voiced speech excitation. However, the exact theoretical convergence analysis is only possible in the case of stationary signals [6].

In the next section, a comparative experimental analysis of the application of CEUC and CIQC algorithms in the robust recursive identification of the time-varying AR speech model is presented.

4. EXPERIMENTAL ANALYSIS

The above - mentioned algorithms are compared according to the results of time-varying AR modeling of a real speech signal. The signal consists of five isolated spoken vowels ("a", "e", "i", "o" and "u") from one speaker. The signal is sampled with f_s=10kHz and preemphasized with q=1. The algorithms are used to identify the AR speech model of the 10th order. The objective quality measure is the MAR (Mean Absolute Residual) criterion: $J = 1/M \cdot \sum_{i=1}^{M} |y(i) - \hat{y}(i)|$, where $y(i)$ is the speech sample at the i-th instance, $\hat{y}(i)$ is its linear prediction, and M is the total number of speech samples. The other quality criteria that are presented in this Section are: bias, variance, adaptiveness of AR parameter estimates obtained using the proposed algorithms, and sensitivity to pitch impulses of the obtained AR parameter estimates.

Table 1 shows means (E) and standard deviations (σ) of the MAR criterion values obtained through the analysis of the five vowels using the proposed robust recursive AR speech analysis procedure with CEUC and CIQC clustering algorithms.

Table 1: Means (E) and standard deviations (σ) of the MAR criterion values obtained in the vowel analysis

Vowel	Length	CEUC		CIQC	
		E	σ	E	σ
A	3690	52.26	2.147	49.58	0.314
E	3690	74.14	0.578	72.08	0.389
I	3690	40.21	0.699	39.78	0.341
O	3690	28.22	0.475	27.88	0.793
U	3690	10.84	0.134	10.81	0.248

The values presented in Table 1 are calculated according to the following ten values of N (length of speech segment): 50, 70, 90, 100, 150, 200, 250, 300, 350, and 400 speech samples.

As another quality criterion, we accepted the comparative methodology presented in [5]. Namely, the estimated trajectories are compared to the reference parameter trajectory obtained by the standard LPC sliding window method with the window length smaller than the estimated pitch period. In the experiments we used the standard LPC covariance method with a sliding window length of NL=40 samples. The tops of this trajectory present the best parameter estimates due to the assumption that the LPC analysis window in these moments includes speech samples from the closed-glottis period [5]. Figures 1, 2, 3, 4, and 5 show the estimated trajectories of the first AR parameter (AR_1) obtained using the proposed robust recursive procedure with the application of CEUC and CIQC clustering algorithms in analyzing the five vowels, respectively.

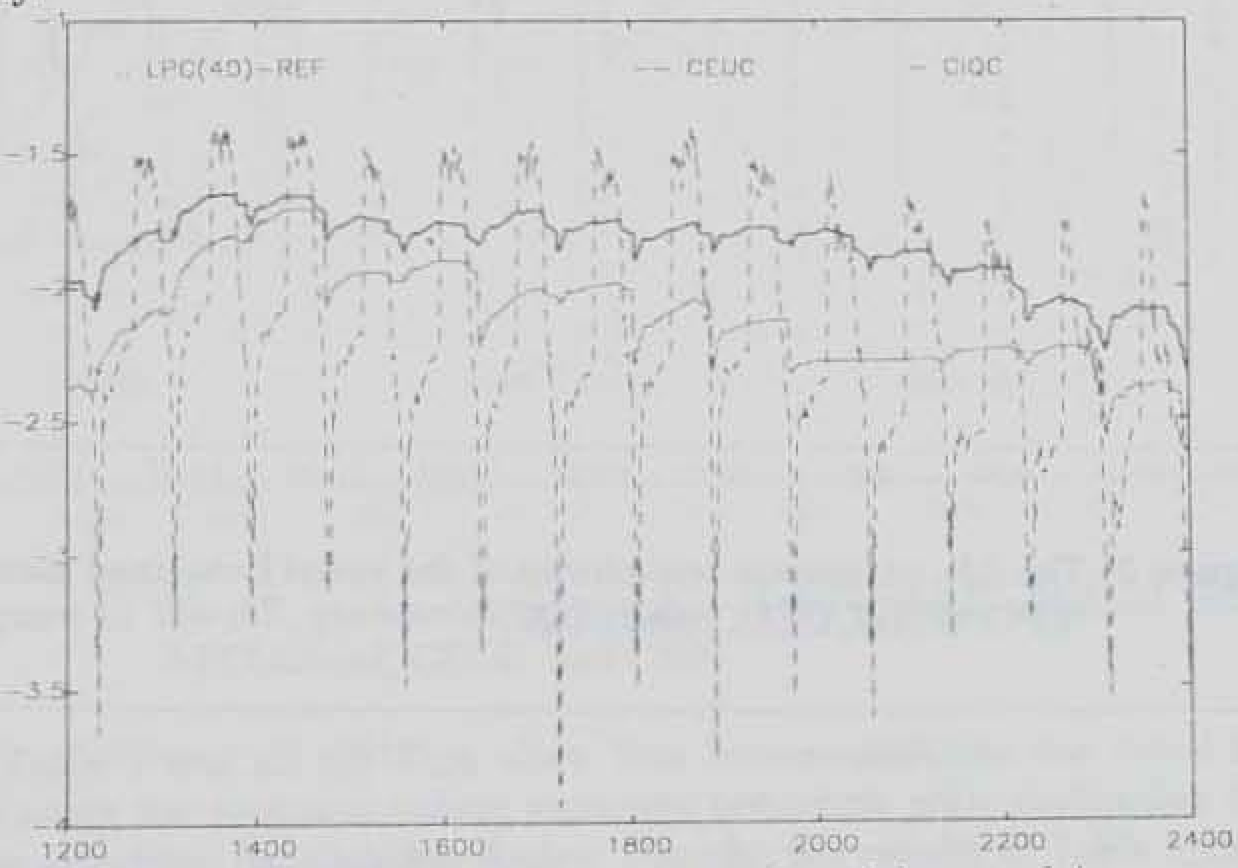

Figure 1: The AR_1 parameter trajectories of the vowel A obtained using: LPC(40)-ref, CEUC, and CIQC.

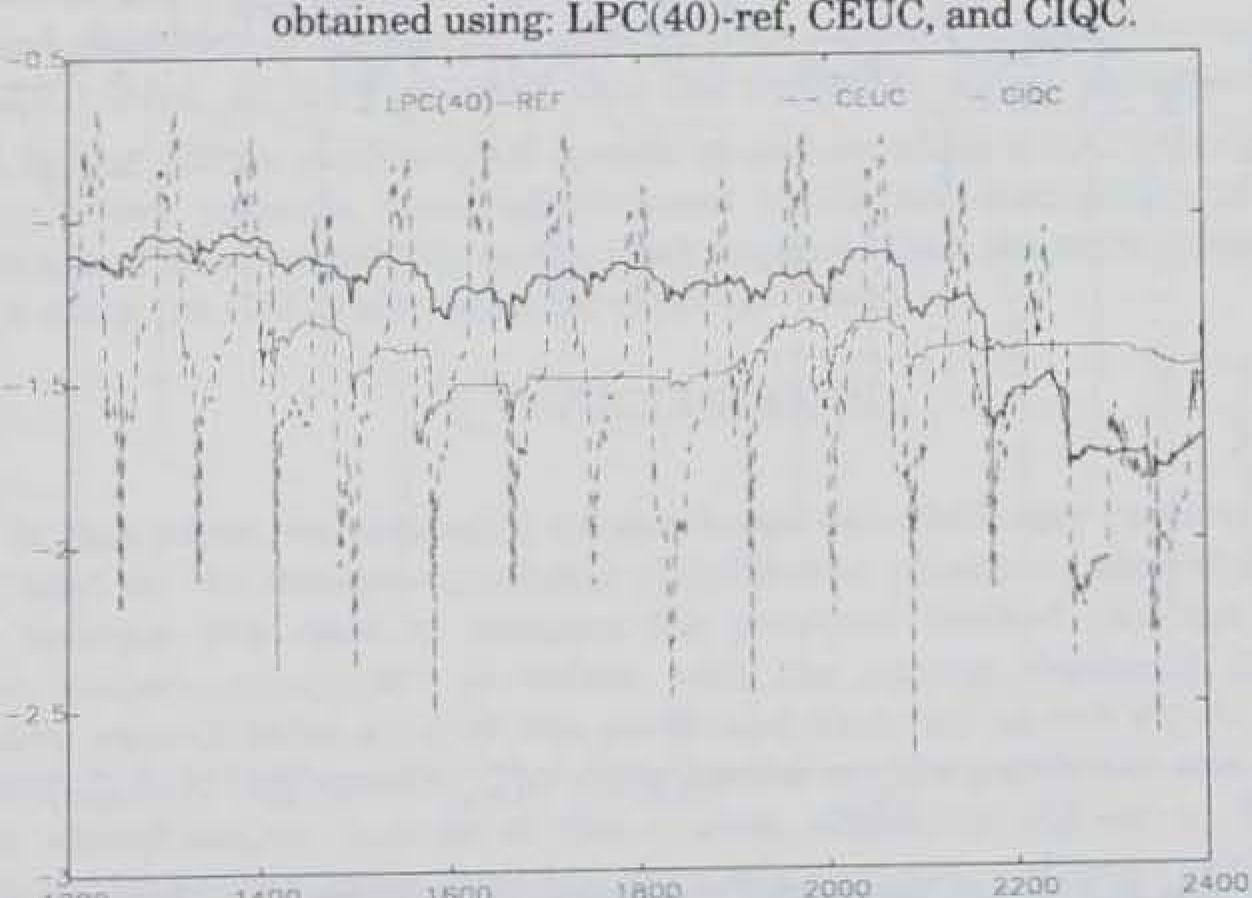

Figure 2: The AR_1 parameter trajectories of the vowel E obtained using: LPC(40)-ref, CEUC, and CIQC.

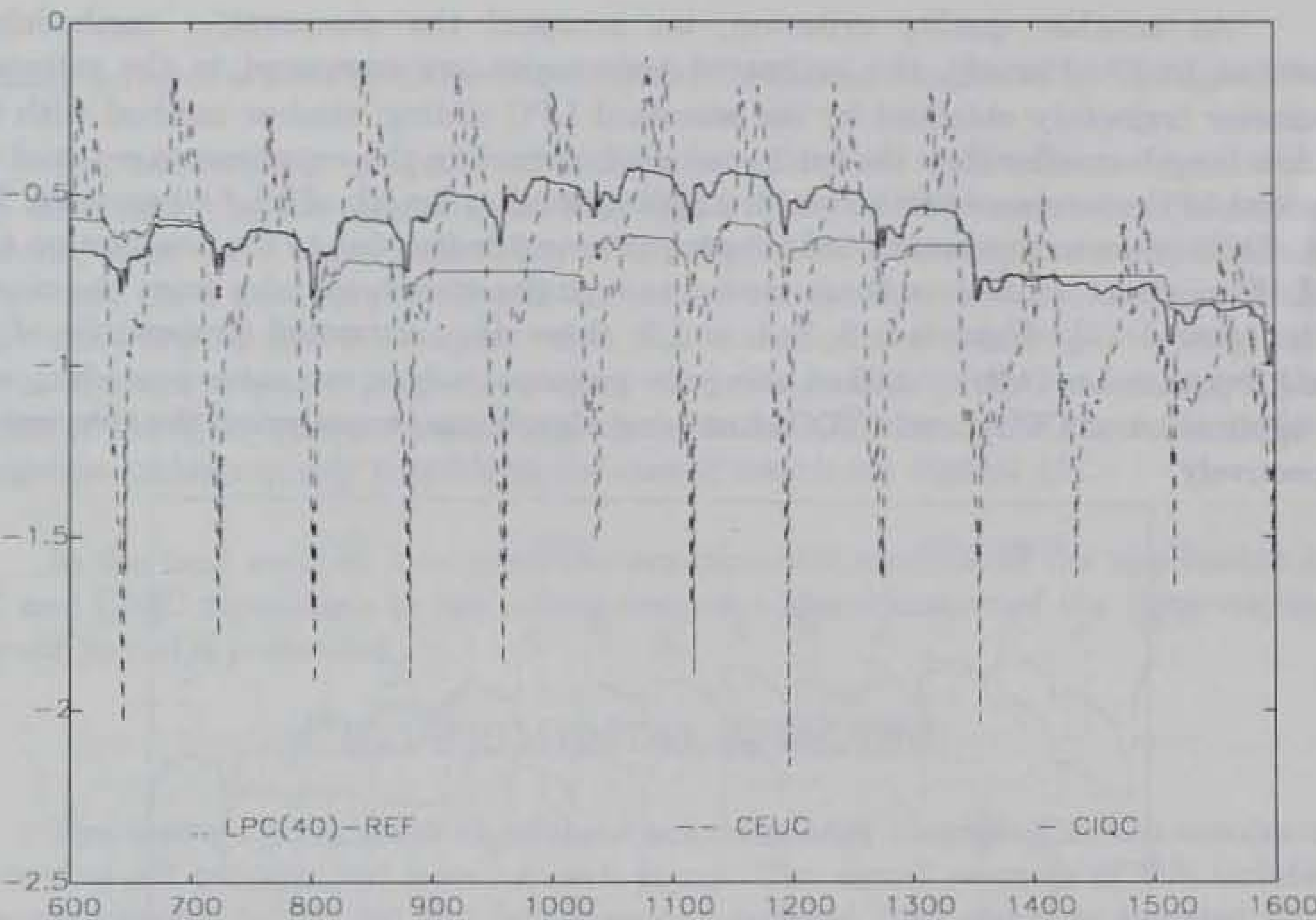

Figure 3: The AR_1 parameter trajectories of the vowel I obtained using: LPC(40)-ref, CEUC, and CIQC.

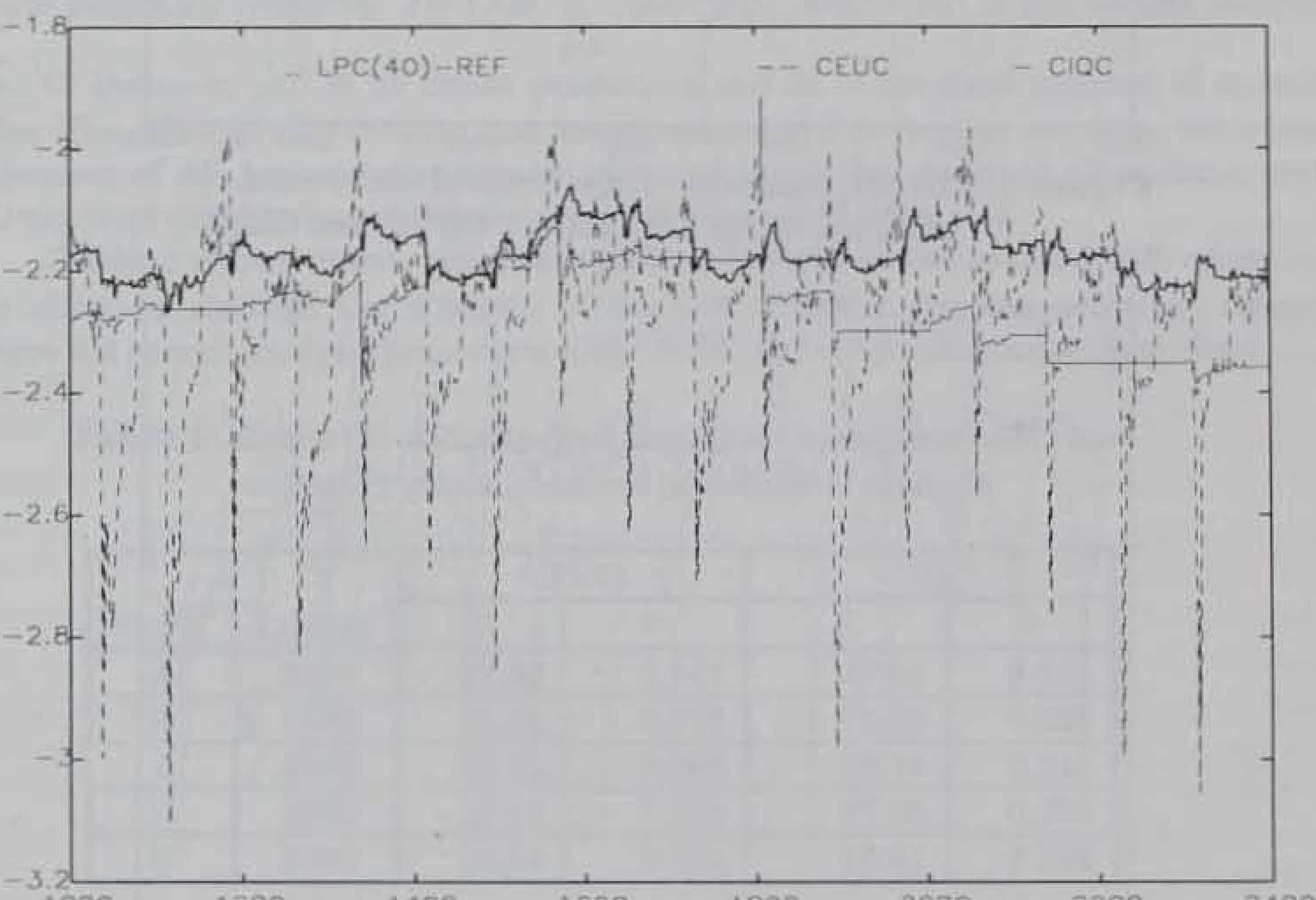

Figure 4: The AR_1 parameter trajectories of the vowel O obtained using: LPC(40)-ref, CEUC, and CIQC.

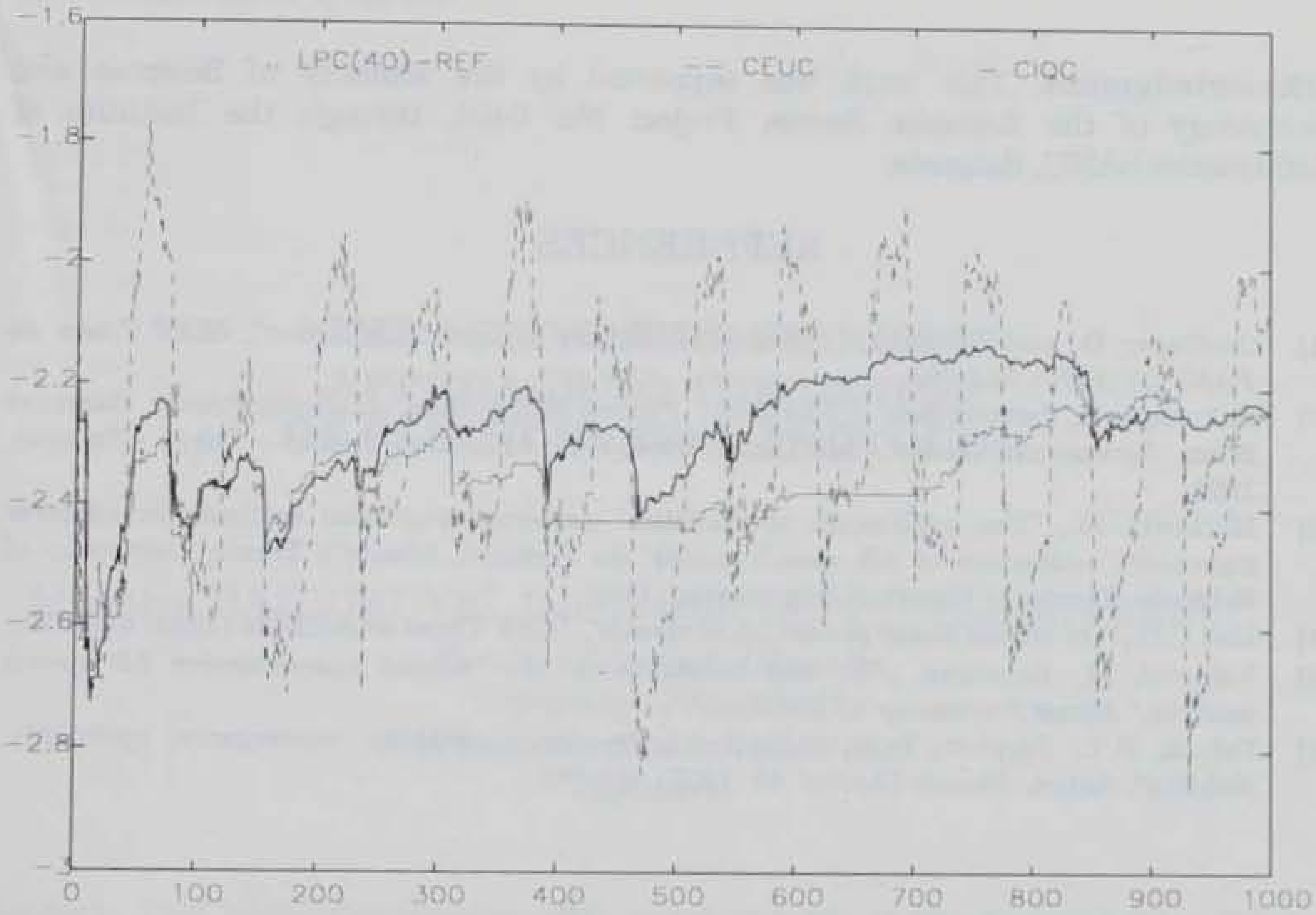

Figure 5: The AR_1 parameter trajectories of the vowel U obtained using: LPC(40)-ref, CEUC, and CIQC.

Table 1 and all the Figs. show that better results in the vowel analysis are obtained using the proposed robust recursive procedure with application of the CIQC clustering algorithm for classifier design. Namely, the results in Table 1 show that the application of the CIQC algorithm produces the smallest objective criterion values (E values) and global lower sensitivity to the value of segment length N (σ values, except for O and U). Also, all the Figs. show that the trajectory of AR_1 parameter estimates obtained by the robust recursive AR speech procedure using the CIQC algorithm has lower bias, lower variance, more adaptiveness to the non-stationarity of the model parameters, and lower sensitivity to the pitch impulses than the same robust recursive procedure using the CEUC algorithm for classifier design.

5. CONCLUSIONS

In this paper, we presented a frame-based non-stationary pattern recognition method based on the iterative quadratic classification clustering algorithm. A speech analysis example was used to compare the proposed method and the same non-stationary pattern recognition procedure with the c-mean clustering algorithm. A comparative experimental analysis was performed on a real speech signal, on isolated spoken vowels from one speaker. The experimental results justify the use of iterative quadratic classifications instead of the c-mean clustering algorithm. It has been observed that lower bias, lower variance, more adaptiveness, and lower pitch sensitivity of the estimated parameter trajectories are obtained by the proposed method.

Acknowledgment. This work was supported by the Ministry of Sciences and Technology of the Republic Serbia, Project No. 0403, through the Institute of Mathematics SANU, Belgrade.

REFERENCES

[1] Gutfinger, D., and Sklansky, J., "Robust classifiers by mixed adaptation", *IEEE Trans. on PAMI* 13 (1991) 552-567.

[2] Fukunaga, K., *Introduction to Statistical Pattern Recognition,* Academic Press., Harcourt Brace Jovanovich, Boston - San Diego - New York - London - Sydney - Tokyo - Toronto, 1990.

[3] Marković, M., "The application of statistical pattern recognition methods in adaptive parameter estimation of AR speech model" (in Serbian), Master's Thesis, University of Belgrade, Faculty of Electrical Engineering, 1992.

[4] Lee, C.H., "On robust linear prediction of speech", *IEEE Trans on ASSP* 36 (1988) 642-650.

[5] Veinović, M., Kovačević , B., and Milosavljević, M., "Robust non-recursive AR speech analysis," *Signal Processing* 37 (1994).

[6] Polyak, B.T., Tsypkin, Ya.Z., "Adaptive estimation algorithms: convergence, optimality, stability", *Autom. Remote Control* 40 (1979) 387-391.