FACTA UNIVERSITATIS (NIŠ)
SERIES: ELECTRONICS AND ENERGETICS VOL. 12, NO. 3 (1999), 103-119

UDC 621.396

QUADRATIC CLASSIFIER WITH HEURISTICALLY DECISION THRESHOLD IN ROBUST AR MODELING OF SPEECH

Milan Ž. Marković

Abstract. Possible applications of quadratic classifier with heuristically decision threshold in robust AR speech analysis and speech compression schemes are considered. A robust recursive procedure based on WRLS algorithm with VFF and the frame-based quadratic classifier with heuristically decision threshold is proposed for identification of nonstationary AR model of speech. Additionally, possible application of the proposed procedure in standard CELP 4800 b/s speech compression algorithm is evaluated. Experimental analysis is done based on the results obtained through analyzing speech signal with voiced and mixed excitation frames. Obtained results show that the proposed robust recursive procedure achieves more accurate AR speech parameter estimation, provides improved tracking performance, and could be used to solve some speech compression tasks more efficiently.

1. Introduction

Linear prediction coding (LPC) of speech signal [6] is based on a linear model of speech production system, given by

$$s(k) + \sum_{i=1}^{p} a_j s(k-i) = e(k) \tag{1}$$

where $s(k)$ is a speech sample, $\{a_i\}$ $(i = 1, \ldots, p)$ are the parameters of AR model (LPC parameters) of order p and $e(k)$ is a sample of speech excitation signal. In the conventional LPC analysis, the LPC parameters are estimated by either autocorrelation or covariance method [6]. Both algorithms minimize the sum of squared residuals (a difference between a speech sample and

Manuscript received Jan. 10, 1999.
The author is with Institute of Applied Mathematics and Electronics, Kneza Miloša 37 11000 Belgrade, Serbia, e-mail: markovic@kiklop.etf.bg.ac.yu .

its linear prediction) representing the least squares (LS) type algorithms. These algorithms are optimal if the excitation signal is an innovation random process of white Gaussian noise type.

However, there are two main problems in application of the conventional LPC methods. The first problem consists of an inherent nonstationarity of speech production system, while the second problem is, in fact, that the speech excitation does not match the assumption of white Gaussian noise, particularly on the voiced speech frames. In the other words, the AR model cannot adequately represents the voiced speech.

In order to solve both of the above mentioned problems, robust recursive procedures with a frame-based statistical pattern recognition approach for efficient identification of nonstationary AR speech model are proposed in [10,7,9]. These algorithms are based on the weighted recursive least squares (WRLS) algorithm with variable forgetting factor (VFF), as an estimation procedure, and a frame-based quadratic classifier of nonstationary signals, as an adaptive classifier. In fact, the iterative quadratic classification method [1] for design of the frame-based classifier is proposed in [10], and its modified version for real-time applications is proposed in [7,9]. It is shown in [7,9] that proposed modification is rather insensitive to an inappropriateness of the assumed classification model and, consequently, it is recommended for the use in the frame-based nonstationary pattern recognition systems. The basis of the mentioned procedures is the assumption of two-class nature of the voiced speech excitation, such that a large part of the excitation is from a normal distribution with a very small variance while a small part is also from the normal distribution but with a much bigger variance. Since the quadratic classifier is the optimal classifier in case of Gaussian distributed data, this approach is emphasized compared to other robust nonrecursive [4,17] and recursive [3] speech analysis approaches based on Huber's M-estimation theory, as well as robust recursive procedure based on Kalman filtering [19].

In order to improve tracking performance of the robust recursive algorithm, described in [7,9], a procedure based on the quadratic classifier with a heuristically decision threshold is proposed. Comparative experimental analysis is done through processing the real speech signal with voiced and mixed excitation frames. Besides, a possible application of the quadratic classifier with heuristically decision threshold in the nonrecursive robust estimation procedure in standard CELP 4.8 kb/s speech coder is evaluated.

The paper is organized as follows. Section 2 is dedicated to a brief description of the proposed recursive procedure and its application to robust

AR speech analysis. A possible application of the proposed nonrecursive robust procedure in the standard CELP 4800 b/s speech coder is considered in Section 3. Comparative experimental analysis is presented in Section 4 while conclusion is given in Section 5.

2. A Novel Robust Recursive AR Speech Analysis Procedure

The equation (1) can be rewritten in the linear regression form

$$s(k) = Z^T(k)\theta + e(k) \tag{2}$$

where $\vec{\theta} = \{a_1 \dots a_p\}$ is the vector of LPC parameters, and $Z^T(k) = \{-s(k-1)\dots - s(k-p)\}$ is the observations vector. Similarly as the nonrecursive sliding window methods [4,17], the application of WRLS algorithm with VFF represents a way for solving the problem of identification of nonstationary AR model of speech production system. Based on the equation (2), the WRLS algorithm with VFF is given by [5]

$$\varGamma(k) = \frac{1}{\rho}\Big[\varGamma(k-1) - \frac{\varGamma(k-1)Z(k)Z^T(k)\varGamma(k-1)}{\rho + Z^T(k)\varGamma(k-1)Z(k)}\Big] \tag{3}$$

$$\theta(k) = \theta(k-1) + \varGamma(k)Z(k)\big[s(k) - Z^T(k)\theta(k-1) \tag{4}$$

where $\varGamma(k)$ is the gain matrix and ρ is the variable forgetting factor (VFF). The value of VFF less than one makes the WRLS algorithm adaptive to the nonstationarity of LPC parameters. In order to obtain the reliable estimates of nonstationary LPC parameters, the value of VFF is determined at each time instances by the modified generalized likelihood ratio (MGLR) algorithm [13], which enables fully automatic detection of the instants of abrupt changes in stationarity of speech signal. In the other words, the value of VFF changes at each time instances according to the amount of LPC parameter variability, which is expressed through the value of the MGLR discrimination function [3]. The MGLR algorithm is based on the consideration of signal samples from three frames (windows): reference, test and joint window, which is a concatenation of the preceding two windows, see Fig. 1. During analysis procedure, all of the three windows "slide" with a one-sample step keeping the fixed length and relationship. For given time instance n, the MGLR discrimination function D could be expressed as follow [13]

$$D(n,N) = L(n-N+1, n+N) - L(n-N+1, n) - L(n+1, n+N) \tag{5}$$

$$L(c,d) = (d-c+1)\ln\left[\frac{1}{d-c+1}\sum_{j=c}^{d} e_j^2\right] \tag{6}$$

where N is a window length, and e_j is a residual signal, i.e. $\hat{e}_j = s(j) + \sum_{i=1}^{p} \hat{a}_i s(j-i)$, with $\hat{a}_i$ being an estimate of a_i. A strategy of choosing the VFF at each time instance is based on the fact that the discrimination function $D(n, N)$ is proportional to the degree of nonstationarity of signals. Consequently, letting ρ_{max} when D_{min} and ρ_{min} when D_{max}, as well as by taking the linear interpolation between these values, is intuitively a good choice [3].

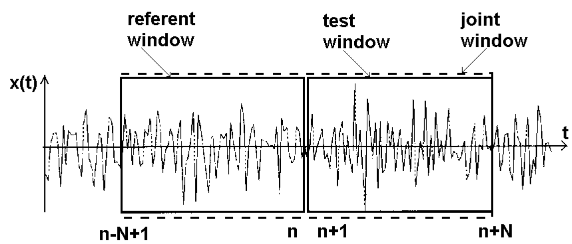

Fig. 1. Three analysis windows in the MGLR algorithm

The relationship between the actual value of ρ and the current value of the MGLR discrimination function is given in Fig. 2. The values of ρ_{max}, ρ_{min}, D_{max}, $_{Dmin}$ must be determined in advance.

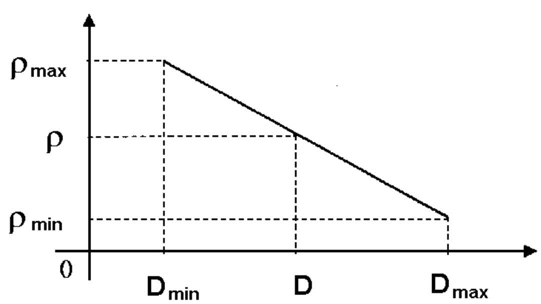

Fig. 2. Relationship between an actual value of ρ and current value of MGLR discrimination function D.

In order to solve the problem of inappropriateness of AR modeling of speech production system, particularly on the voiced frames, a procedure for robustifying the WRLS algorithm with VFF by applying nonstationary pattern recognition method is proposed in [10,7,9]. This procedure consists of the application of frame-based quadratic classifier in a combined nonrobust/robust recursive AR speech analysis scheme. The nonrobust procedure represents the WRLS algorithm with VFF ($\rho < 1$) given by the equations (3) and (4). In this case, as mentioned before, ρ changes at each time steps according to the value of MGLR discrimination function, as proposed in [3]. On the other hand, the robust procedure is the WRLS algorithm with variable factor $\rho > 1$, which changes according to the value of corresponding residual sample. In this case, the value of ρ is heuristically determined by the expression: $\rho = 1/(1 - |r_{norm}|/2)$, where r_{norm} is the normalized value obtained by dividing the current residual with the maximal residual on the current frame. In addition, the maximal residual value is updated on the frame-by-frame manner. In this way, the algorithm assigns less weight to the large residuals, so that it is robust in the sense of its insensitivity to the spiky excitations on the voiced frames.

In this heuristic procedure, the frame-based quadratic classifier of nonstationary signals is used to classify the residual speech samples into the two classes. The first class consists of "small" residual samples and the second one consists of "large" residual samples. Fig. 3 represents the example of residual signal and two corresponding classes, obtained after applying the proposed classification procedure on the natural speech signal with voiced speech frames (Serbian vowel "I"). The classification of the k-th residual sample selects either the nonrobust (first class) or the robust (second class) recursive AR procedure for LPC parameter estimation at the k-th time instance. In this case, the classifier is very simple, i.e. it is one-dimensional, and the mean vectors and covariance matrices are reduced to the means and variances, respectively. The classification consists of two steps: *initialization* and *adaptation*.

Initialization: On the initial frame of signal one has to determine the following: (1) the starting LPC parameter vector which is used as the initial condition for the proposed recursive procedure; (2) the initial maximal residual value; and (3) the initial partition of the frame. The starting LPC parameters are obtained by applying the conventional nonrecursive covariance LPC method [6] on the initial frame (good results are obtained with the initial frame length of 100 samples). The initial quadratic classifier is obtained applying the iterative quadratic classification procedure [10], based on the initial partition. Having calculated the initial maximal residual value, this partition is obtained by comparing the obtained normalized residuals with the threshold of value 0.5 (the residuals less than the threshold are assigned to the first class, otherwise they are assigned to the second class). As well as the c-mean algorithm, the iterative quadratic classifications clustering algorithm is derived from the general clustering algorithm, described in [1]. Assume that we want to classify the N samples, $X_1, \dots, X_N$ into the

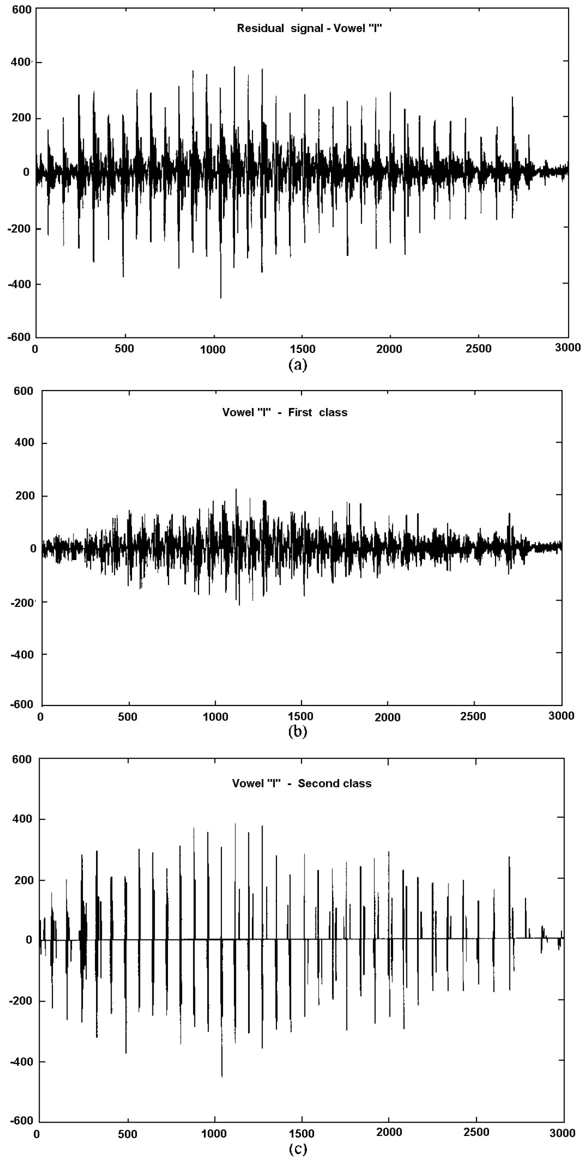

Fig. 3. An example of the nonstationary quadratic classifier application in robust recursive AR speech analysis Serbian vowel "I": (a) Residual signal; (b) Samples classified in the first class; (c) Samples from the second class.

one of L classes, $\omega_1, \ldots, \omega_L$, where L is assumed to be given. The iterative quadratic classifications clustering algorithm has the form:

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

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 ω_i is estimated by the ratio of the number of samples within the class ω_i and the total number of samples.

3. If the classification of any X_j is changed, calculate the $P_i(l+1)$, $M_i(l+1)$, and $\Sigma_i(l+1)$ for the new class assignment, and go to Step (2). Otherwise, stop.

Adaptation: The initial classifier is then applied for classifying the residual speech samples obtained by the proposed recursive AR speech analysis on the next frame of signal with size N. The result of the k-th residual sample classification invokes either the nonrobust recursive procedure ($\rho < 1$) or robust one ($\rho > 1$) to estimate the vector of LPC parameters in the k-th time instance. The obtained parameter vector is used to determine the $(k+1)$-th residual sample and the procedure is continuing. The classification result of entire frame represents the initial partition of that frame and is used to produce the initial quadratic classifier for the next frame of size N, without the use of the iterative procedure (RTQC algorithm [7,9]), and so on.

In this paper, a modification of the RTQC algorithm (named RTQCH algorithm) which is based on the quadratic classifier with heuristically decision threshold is considered. Namely, as it is shown in [1], the use of the optimal decision threshold in Bayes error estimation procedure when the limited training data set is available is highly inefficient, and a heuristically determination of the decision threshold is suggested. The problem of the limited training data set is also presented in the application of the quadratic classifier in AR speech parameter estimation and, thus, a similar procedure for determination of the decision threshold is proposed. The value: $Thresh = (\sigma_1 + \sigma_2)/2$, is adopted as the heuristically threshold where σ_1, σ_2 represent standard deviations of the first and second class, respectively.

The convergence property of the proposed robust estimation algorithm is mainly determined by the conventional WRLS algorithm with VFF, since the a priori probability of the first class is significantly greater 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 WRLS 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 signal [16].

3. Robust Nonrecursive LPC Parameter Estimation in Standard Celp 4.8 Kb/s

This Section deals with the partial problem of CELP coding algorithms: the estimation of LPC spectral parameters. In the standard CELP 4800 b/s coder, USFS1016 CELP 4800 b/s [18], the autocorrelation method with 30 ms Hamming window is adopted as the LPC parameters estimation procedure, without presenting experimental comparison with some other LPC methods. In this sense, a comparative analysis of the influence of the LSP parameters quantization and interpolation, used in [18], to the spectral characteristics of four standard LPC methods (autocorrelation with Hamming window, covariance, modified covariance, and lattice method) is presented in [12]. The comparative experimental analysis was done based on the three different spectral measures related to the RMS LOG spectral measure: likelihood ratios, cosh measure and cepstral distance [2]. The experimental results, presented in [12], have not justified the use of the autocorrelation method with Hamming window in USFS1016 CELP 4800 b/s speech coder. Namely, the best results, e.g. the smallest spectral degradation between the unquantized and quantized interpolated LSP parameters, were obtained by using the modified covariance method.

A possibility of using robust method based on Huber's M-estimation theory (RLPC method) for efficient LPC parameters estimation in standard CELP 4800 b/s coder is elaborated in [14]. Namely, the comparative analysis of the influence of the quantization and interpolation of LSP parameters to the spectral characteristics of the RLPC and standard LPC parameter estimation methods was presented. The experimental results, presented in [14], justified the use of the proposed RLPC estimation procedure in USFS1016 speech coding algorithm. However, the Huber's robust estimation procedure is iterative in its nature, and could make a considerable increasing of the overall complexity of the standard CELP algorithm. To overcome this problem, the use of a heuristic two-stage sample-selective LPC method, which is slightly more complex than the autocorrelation method, is proposed in [8,11]. The application of this method (which is a modified version of the corresponding method proposed in [15]) in the USFS1016 4.8 kb/s results in smaller LPC spectral degradation, compared to the standard LPC methods. In the other words, the experimental results, presented in [8,11], justified the use of the heuristic robust sample-selective LPC procedures in the USFS1016 algorithm. In order to further decrease the LPC spectral degradation in USFS1016 algorithm, an application of the frame-based quadratic classifier of nonstationary signals in the mentioned heuristic sample-selective LPC procedure is considered.

The proposed sample-selective LPC method is realized as a two-stage standard LPC method and represents the modification of the method developed in [15]. Namely, instead of excluding the whole prediction equations that lead to the very large residual values, as proposed in [15], we have adopted zero values in the observation matrix for the speech samples corresponding to the very large residual values. These large residual values are selected by using the frame- based quadratic classifier of nonstationary signals. The proposed robust sample-selective LPC parameter estimation procedures based on the quadratic classifier consist of the following steps:

1. Applying one of the standard LPC methods on the windowed speech frame (the first stage of the procedure).
2. Inverse filtering of the speech on the given frame by using the estimated inverse filter to obtain the residual signal.
3. Calculating the maximum residual value and normalizing the residual signal on the entire frame.
4. Applying the given or another standard LPC method (the second stage of the procedure) with the modified observation matrix of the given speech frame. Quadratic classifier is used to classify the normalized residual samples into the two classes, the class consisting of "small" residual samples and the class consisting of "large" ones. Namely, in case of the speech samples for which the corresponding normalized residual values are classified into the class of the "large" residual values by using the above described adaptive real-time frame-based quadratic classifier, zero values are adopted in the modified observation matrix. In this paper, we consider the use of the quadratic classifier with optimal decision threshold [7,9], and with a heuristically threshold, obtained by the procedure similar to the one proposed in [1]. The value: $Thresh = (\sigma_1 + \sigma_2)/2$, is also adopted as the heuristically threshold in this case.

3. Experimental Analysis

The efficiency of the proposed algorithm is tested on natural speech signal. The signal consists of five isolately spoken serbian vowels ("A", "E", "I", "O", "U") and ten isolately spoken digits ("1", "2", ... ,"0") from one speaker. The signal is sampled with $f_s = 10$ kHz and preemphasized with

$q = 1$ [17]. All experimental results are obtained by using the 10-th order AR model. As the objective quality measure, the MAR (Mean Absolute Residual) criterion is used [7]

$$J = \frac{1}{M} \sum_{i=1}^{M} |s(i) - \hat{s}(i)| \tag{7}$$

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 speech samples. In the all experiments the following bounds of VFF are used: $\rho_{max} = 0.99$, and $\rho_{min} = 0.95$, as proposed in [3]. In the case of the MGLR discrimination function bounds, in the all experiments, we used $D_{min} = 0$, while D_{max} was initially set to 200 and then updated during analysis of the given speech frame [3]. Table 1 shows the MAR criterion values obtained through the analysis of five vowels and ten digits by using the WRLS algorithm with VFF and the proposed robust recursive AR speech analysis procedure based on the quadratic classifier with the optimal (RTQC) and heuristically decision threshold (RTQCH). The MAR values for robust procedures, presented in Table 1, are obtained by using the $N = 100$ speech samples as an adaptation frame. Also, numbers of residual speech samples classified in the first (CLASS1) and second (CLASS2) class are presented in the Table 1.

As the other evaluation criteria, bias, variance, and sensitivity to the intensity of pitch impulses of the estimated AR speech parameters are used. Fig. 4, and 5 show the typical examples of estimated trajectories of the first LPC parameter (AR1), obtained by the WRLS algorithm with VFF and the two versions of the proposed robust recursive procedure (RTQC and RTQCH) for the digits "2" and "8", respectively. As for the RTQC and RTQCH algorithms, the trajectories, shown on Fig. 4 and 5, are obtained by using the speech frame length (learning data set length) of $N = 100$ speech samples. Estimated AR1 parameter trajectories are compared in terms of the "reference trajectories" [17] of estimated LPC parameters, which were obtained using a rectangular sliding window shorter than the pitch period and the nonrecursive covariance LPC method. Namely, the experimental analysis has shown that the LPC parameter estimates are most accurate when the data sliding window is shorter than the pitch period, comprising almost a whole closed-glottis interval but not the instant of the glottal closure (a such part of signal corresponds quite well to the linear AR model with white Gaussian excitation and, consequently, the given LPC method gives the optimal estimates) [17]. In this sense, local maximum tops of the LPC-REF AR1 parameter estimates trajectories (see Fig. 4 and 5) represent

Table 1. MAR criterion values: RTQC and RTQCH algorithms.

Test Signal	WRLS with VFF	RTQC			RTQCH		
		MAR	CLASS1	ClASS2	MAR	CLASS1	CLASS2
A	52.34	50.42	3271	419	50.41	3269	421
E	74.56	73.05	3459	231	73.01	3332	358
I	39.72	39.81	3354	336	39.57	3264	426
O	27.87	27.00	3521	169	27.13	3311	379
U	10.40	10.32	3552	138	10.29	3266	424
1	36.94	34.11	6157	533	34.97	5904	786
2	28.56	27.36	6217	473	27.11	5598	1092
3	28.78	26.84	5221	469	28.48	4536	1154
4	27.07	24.46	5848	842	24.51	5527	1163
5	19.39	17.34	6099	591	17.26	5511	1179
6	22.28	21.12	6779	911	21.58	5566	2124
7	34.99	32.90	6158	532	32.65	5634	1056
8	17.89	18.82	6432	1258	17.42	6207	1483
9	38.72	37.44	5110	580	37.52	4646	1044
0	21.96	21.23	5222	468	21.32	4257	1433

the points of the "reference trajectories", i.e. the best parameter estimates. Based on the experimental results, presented in Table 1, and Fig. 4 and 5, it can be concluded that the trajectories of AR1 parameter estimates, obtained by the proposed robust recursive AR speech analysis procedure based on the quadratic classifier with heuristically decision threshold (RTQCH) possess improved tracking characteristics, lower bias, lower variance, and lower sensitivity to the intensity of pitch impulses than the conventional WRLS with VFF, and earlier version of the robust procedure (RTQC algorithm). The main advantage of the RTQCH algorithm is in more robustness (larger values in the column CLASS2 of the Table 1) compared to the RTQC algorithm.

As for evaluation of the proposed robust procedure in speech coding domain, a possible application of the RTQC and RTQCH algorithms for nonrecursive AR speech parameter estimation in USFS1016 4.8 kb/s CELP speech coder [18] is analyzed. The test speech base consists of 4 files, each containing 10 isolated digits, spoken by three male speakers (R1, R4, and R9) and a female speaker (R14). File lengths of R1, R4, R9, and R14 are 64000, 68000, 72000, and 77000 samples, respectively. LPC analysis of the 10th order is applied to the nonpreemphasized speech signal on the nonoverlapping frames of 30 ms, as in [18]. The standard and robust LPC

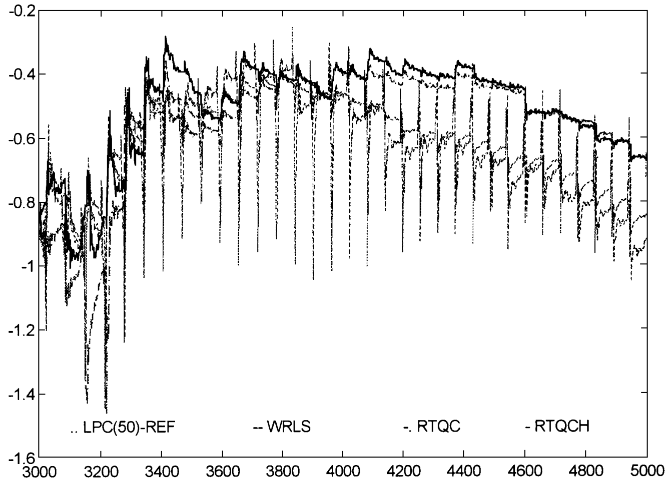

Fig. 4. Trajectories of AR_1 parameter estimates obtained by using: LPC(50)-REF, WRLS with VFF, RTQC, and RTQCH algorithms, in analyzing of digit: "2".

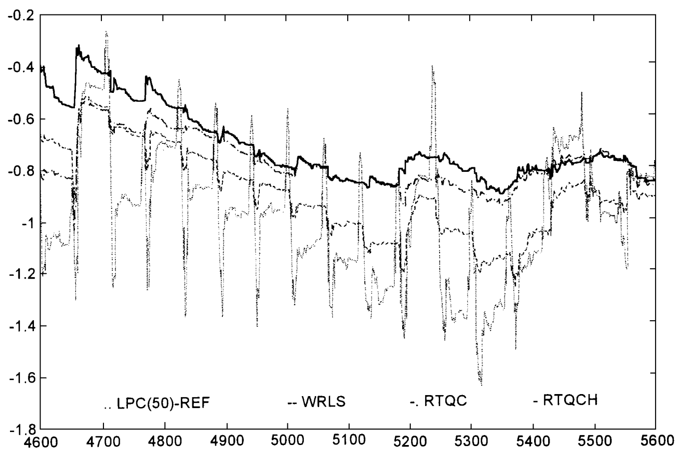

Fig. 5. Trajectories of AR_1 parameter estimates obtained by using: LPC(50)-REF, WRLS with VFF, RTQC, and RTQCH algorithms, in analyzing of digit: "8".

methods are performed in such a way that LPC parameters, obtained by using the proposed methods (with 15 Hz bandwidth expansion), are first transformed into LSP parameters, quantized and then interpolated, in order to determine LSP values to be used for each subframe. Spectral distortion measures (denoted by "Q") are computed between the signal spectra on the given speech frame, obtained by corresponding unquantized and quantized LSP parameters. The same measures (denoted by "Q+I") are computed on the subframe level. As the objective spectral criterion, the cepstral distance [2], as the spectral measure related to the RMS LOG spectral measure, is used in all the experiments. RMS LOG spectral measure is given by

$$(d_2)^2 = \int_{-\pi}^{\pi} |V(\theta)|^2 \frac{d\theta}{2\pi} \tag{8}$$

where:

$$V(\theta) = \ln\left[\frac{\sigma^2}{|A(e^{j\theta})|^2}\right] - \ln\left[\frac{(\sigma')^2}{|A'(e^{j\theta})|^2}\right] \tag{9}$$

The Fourier series expansion for the model log spectrum can be written as

$$\ln\left[\frac{\sigma^2}{|A(e^{j\theta})|^2}\right] = \sum_{k=-\infty}^{\infty} c_k e^{-jk\theta} \tag{10}$$

where $c_0 = \ln[\sigma^2]$, and $c_{-k} = c_k$, are cepstral coefficients. A similar expression holds for the expansion of: $\ln[(\sigma')^2/|A'(e^{j\theta})|^2]$, yielding coefficients $\{c'_k\}$. Applying the Parseval's relation to the $(d_2)^2$ we could define the cepstral distance $u(L)$ as [2]

$$u^2(L) = \sum_{k=-L}^{L} (c_k - c'_k)^2 = (c_0 - c'_0) + 2\sum_{k=1}^{L}(c_k - c'_k)^2 \tag{11}$$

In this paper, we use the $L = 4p$ (p is the filter order). Besides, only the results for the case that the gain constants (σ and σ') are identical and equal to 1 (DM(7) from [2]) are used. Also, the factor 4.34294418 (representing the quotient: $10/\ln(10)$) is used, providing that the values are expressed in decibels. The values of the cepstral distance spectral measure for the test speech base obtained by using the standard [12], Huber's robust [14], sample-selective [8], combined sample-selective LPC methods [11], and robust sample-selective LPC methods based on quadratic classifier with optimal (RTQC) and heuristically (RTQCH) decision threshold, proposed in this paper, are shown in Table 2. Huber's robust LPC method is denoted as RLPC, while the sample-selective LPC methods are denoted as: A-A (two- stage autocorrelation method with Hamming window), C-C (two-stage covariance), M-M (two-stage modified covariance) and L-L (two-stage

Table 2. Summary values of the cepstral distance for the standard and robust LPC methods in USFS1016 CELP 4800 b/s algorithm

Standard/		R1		R4		R9		R14	
Robust	Method	Q	Q+I	Q	Q+I	Q	Q+I	Q	Q+I
Standard	A	490.4	2248.6	509.4	2379.4	606.3	2460.3	483.3	2689.6
	C	492.2	2166.7	505.9	2256.8	618.0	2342.7	487.5	2566.3
	M	488.6	2164.2	506.3	2254.3	619.2	2337.7	488.3	2559.8
	L	487.9	2167.0	506.6	2260.4	618.5	2358.2	487.8	2573.8
Robust	RLPC	448.9	2179.8	462.3	2242.8	550.9	2288.3	477.6	2526.5
Sample-Selective LPC	A-A	396.5	2057.5	405.7	2191.4	431.4	2226.3	435.0	2378.9
	C-C	381.9	1765.4	392.4	1896.5	421.1	1888.8	433.7	2058.9
	M-M	383.1	1759.4	391.5	1887.7	433.4	1881.1	436.4	2051.3
	L-L	392.4	1792.5	399.6	1925.2	431.1	1940.2	435.1	2083.3
Combined Sample Selective LPC	C-A	375.8	1716.1	393.3	1860.6	430.3	1832.4	439.4	2017.3
	M-A	379.0	1726.5	393.4	1869.2	432.0	1830.3	436.4	2021.8
	L-A	380.2	1718.2	392.0	1866.2	433.2	1845.8	436.2	2023.7
	C-M	380.0	1749.3	391.0	1879.4	431.7	1881.9	437.3	2042.7
Sample-Selective RTQC	A-A	397.6	2111.2	419.6	2404.7	441.8	2316.1	432.8	2629.0
	C-C	391.9	1816.3	399.0	2089.7	430.8	1995.4	436.8	2290.5
	M-M	391.6	1803.7	399.0	2097.9	443.9	1979.6	433.9	2290.8
	L-L	399.9	1838.1	403.5	2136.3	432.8	2048.9	430.8	2299.2
Combined Sample-Selective RTQC	C-A	386.4	1768.3	399.1	2040.3	432.5	1905.9	434.2	2227.7
	M-A	386.7	1770.0	399.2	2073.3	438.4	1909.6	432.2	2246.1
	L-A	385.5	1763.6	398.7	2066.3	431.8	1934.0	429.2	2215.0
	C-M	391.7	1802.2	397.8	2063.6	438.3	1976.0	436.5	2269.5
Sample-Selective RTQCH	A-A	381.8	1975.7	407.6	2081.6	426.6	2162.7	429.1	2252.3
	C-C	373.8	1711.7	394.1	1856.3	427.0	1946.5	425.4	1998.4
	M-M	373.6	1685.6	392.4	1859.3	436.0	1892.6	425.8	2073.4
	L-L	378.7	1708.9	390.8	1808.7	437.0	1964.7	426.8	1938.7
Combined Sample-Selective RTQCH	C-A	370.0	1642.7	387.4	1782.1	425.4	1789.7	423.3	1950.9
	M-A	369.7	1644.5	388.3	1806.8	427.8	1800.0	418.6	2001.5
	L-A	370.4	1626.1	390.8	1808.7	431.4	1816.3	426.8	1938.7
	C-M	372.3	1678.8	393.7	1842.1	435.8	1903.9	426.9	2013.1

lattice method). Based on an extensive experimental analysis, we have used the following combinations of the standard LPC methods for the combined sample-selective LPC methods: C-A, M-A, L- A, and C-M.

Based on the experimental results (see Table 2), it can be concluded that the influence of the global LPC spectral representation to the speech signal spectra obtained by the standard LPC and robust methods is smaller for the robust procedures. Furthermore, the comparison between the robust procedures (Huber's, sample-selective, combined sample- selective LPC methods, and corresponding methods based on quadratic classifier) have shown that the smallest influence of the global LPC spectral representation in the USFS1016 algorithm is obtained for the combined sample- selective LPC methods based on quadratic classifier with heuristically determined decision threshold (RTQCH algorithm). Additionally, the best results (see Table 2) are obtained by the two-stage combined sample-selective methods based on quadratic classifier with the covariance method in the first stage and the autocorrelation method with the Hamming window in the second stage of the procedure.

Based on the entire analysis, the proposed robust recursive estimation procedure based on the quadratic classifier with heuristically decision threshold (RTQCH algorithm) is recommended as a possible efficient solution to the problem of AR speech parameter estimation in the speech analysis tasks. Additionally, the possible application of the nonrecursive sample-selective method based on the RTQCH algorithm for the frame-based quadratic classifier design as the robust LPC parameters estimation procedure in standard CELP 4800 b/s speech compression scheme is also justified and thus recommended.

5. Conclusion

In the paper, a new robust recursive procedure for parameter estimation of nonstationary AR model of speech production system based on the WRLS algorithm with VFF and quadratic classifier with heuristically decision threshold is introduced. The comparative experimental analysis of the proposed method, nonrobust WRLS algorithm with VFF, and earlier version of the robust procedure, is performed on the natural speech signal, isolately spoken vowels and digits. Experimental results justify that the method proposed in the paper represents a way to cope with the two main problems of LPC speech analysis, the nonstationarity of LPC parameters and limited validity of AR model of speech, particularly on the voiced frames. Namely, it has been observed that the proposed algorithm is more efficient (lower bias, lower variance, and lower sensitivity to the pitch impulses) compared to the conventional WRLS algorithm with VFF, and the same robust procedure based on the quadratic classifier with optimal decision threshold. Additionally, the possible application of the quadratic classifier with heuristically decision threshold is experimentally justified in the proposed robust nonrecursive LPC parameters estimation procedure in the standard CELP speech coder. Based on the entire experimental analysis, the proposed robust recursive and nonrecursive procedures based on the quadratic classifier with heuristically decision threshold are recommended as possible solutions for parameter estimation of AR speech model in solving some AR speech analysis and speech compression tasks, respectively.

REFERENCES

1. K.Fukunaga: *Introduction to Statistical Pattern Recognition.* ACADEMIC PRESS INC., New York, 1990.
2. A.H.Gray, Jr., J.D.Markel: *Distance Measures for Speech Processing.* IEEE Trans. of ASSP, Vol. ASSP-24, No. 5, Oct. 1976., pp. 380-391.
3. B.Kovacevic, M.Milosavljevic, M.Veinovic: *Robust recursive AR speech analysis.* Signal Processing, Vol. 44, 1995, pp. 125-138.
4. C.-H.Lee: *On robust Linear Prediction of Speech.* IEEE Trans. on ASSP, Vol. 36, No. 5, May 1988., pp 642-650.
5. L.Ljung: *System Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, NJ, 1987.
6. J.D.Markel, A.-H.Gray, Jr.: *Linear Prediction of Speech.* New York: Springer-Verlag, 1976.
7. M.Markovic: *The application of statistical pattern recognition methods in adaptive parameter estimation of AR speech model.* (in serbian) Master Thesis, University of Belgrade, Faculty of Electrical Eng., 1992.
8. M.Markovic: *The application of sample-selective LPC method in standard CELP 4800 b/s speech coder.* Proc. of the Third IEEE International Conference on Electronics, Circuits, and Systems, ICECS'96, October 13-16, Rodos, Greece, Vol. 1, pp. 506-509.
9. M.Markovic, B.Kovacevic, M.Milosavljevic: *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, Nov. 1996., pp. 456-460.
10. M.Markovic, M.Milosavljevic, B.Kovacevic: *Clustering in Non-Stationary Pattern Recognition Systems.* SIGNAL PROCESSING VII: Theories and Applications, M. Holt, C. Cowan, P. Grant, W. Sandham (Eds.), 1994. European Assoc. for Signal Processing (Proc. of EUSIPCO-94, Edinburgh, Scotland, 13-16 Sept., 1994), pp. 1803- 1806.
11. M.Markovic, M.Milosavljevic, B.Kovacevic, M.Veinovic: *On Robust LPC Parameter Estimation in Standard CELP 4800 b/s Speech Coder.* IEE Proceedings - Vision, Image and Signal Processing, Vol. 145, No. 1, Feb. 1998, pp. 19-22.
12. M.Markovic, M.Obradovic: *Towards Optimal LPC Spectral Parameters Estimation in Standard CELP 4800 b/s Speech Coder.* Proc. of the 7th IEEE Digital Signal Processing Workshop, Loen, Norway, Sept. 1-4, pp. 161-164.
13. M.Milosavljevic, I.Konvalinka: *The modified generalized likelihood ratio algorithm (MGLR) for automatic detection of abrupt changes in stationarity of signals.* Proc. of the Twenty-Second Annual Conference on Information Science and Systems, Princeton, NY, 1988.

14. M.Milosavljevic, M.Markovic, B.Kovacevic, M.Veinovic: *Robust LPC Parameter Estimation in Standard CELP 4800 b/s Speech Coder.* Proc. of IEEE Region 10 Conference of Digital Signal Processing Applications, TENCON'96, Perth, Western Australia, 27-29 Nov. 1996, pp. 200-203.

15. Z.Miyoshi, K.Yamato, R.Mizoguchi, M.Yanagida, O.Kakusho: *Analysis of Speech Signals of Short Pitch Period by a Sample-Selective Linear Prediction.* IEEE Trans. on ASSP, Vol. ASSP-35, No. 9, September 1987., pp. 1233- 1240.

16. B.T.Polyak, Ya.Z.Tsypkin: *Adaptive estimation algorithms: convergence, optimality, stability.* Autom. Remote Control, Vol. 40, 1979, pp. 387-391.

17. M.Veinovic, B.Kovacevic, M.Milosavljevic: *Robust Non-recursive AR Speech Analysis.* Signal Processing, Vol. 37, No. 2, May 1994., pp. 189-201.

18. USA Federal Standard 1016: *Telecommunications: Analog to digital conversion of radio voice by 4800 bit/second Code Excited Linear Prediction (CELP).* Feb. 14, 1991.

19. T.Yang, J.-H.Lee, K.-Y.Lee, K.-M.Sung: *On robust Kalman filtering with forgetting factor for sequential speech analysis.* Signal Processing, Vol. 63, 1997, pp. 151-156.