COMBINED FILTERING AND RECOGNITION OF THE TYPE OF SEGMENT OF EXCITATION OF NOISY SPEECH SIGNALS

V. I. Gupal, S. Ya. Zhuk, and Yu. A. Mamonov

Izvestiya VUZ. Radioelektronika,
Vol. 35, No. 11, pp. 11-17, 1992

UDC 621.391

Algorithms for filtering and recognition of the type of segment of excitation of noisy speech signals were synthesized on the basis of the apparatus of mixed Markov processes in discrete time. The algorithms obtained are analyzed for the example of processing a real speech signal on a computer.

Methods of suppressing (reducing) distorting noise occupy an important place among methods of processing speech signals. Such processing is important both when solving various speech recognition problems and when compressing speech signals for transmission over digital communication channels. In both cases the effectiveness of processing substantially depends on whether or not the type of segment of stationarity of the speech signal being processed, which in turn is determined by the type of signal of the vocal tract being excited, is known. Adaptive algorithms for combined filtering and recognition of the type of segment of excitation of noisy speech signals, which can be used before further specific processing for the purpose both of compression and recognition, are synthesized in the article on the basis of the apparatus of mixed Markov stochastic processes (MSPs) in discrete time.

A vector model of linear prediction of the following form is used for describing a speech signal with consideration of its nonstationary properties

$$x(k) = F_j(k, k-1)\, x(k-1) + G_j(k)\, w(k), \tag{1}$$

where x(k) - n is the vector of the speech signal being processed; w(k) - δ is a correlated sequence of Gaussian vectors of dimension n × 1 N [0, 1]; $F_j(k, k-1)$, $G_j(k)$ - n × n are matrix coefficients determined from the statistical characteristics of various stationary segments of the speech signal being processed; $j = \overline{1, M}$ is a parameter determining the belonging of the current vector of a fragment to a specific segment of stationarity of the speech signal; M is the number of different segments; x(0) is the initial vector of state N $[\hat{x}(0), \hat{P}(0)]$. The model with random structure (1) permits examining a speech signal in the form of intervals of stationary stochastic sequences. The parameter $\gamma(k) = j$, $j = \overline{1, M}$, which belongs to the class of Markov chains with transition matrix $\Pi_{ij}(k, k-1)$ and initial probabilities $p_i(0)$, $i = \overline{1, M}$, is used for describing the process of replacement of various segments of the speech signal at random times. Four types of segments of stationarity of the speech signal are used in the work: pause (j = 1); vocalized sounds (j = 2), nonvocalized sounds (j = 3), and sounds with a mixed type of excitation (j = 4).

The values of $F_j(k, k-1)$ and $G_j(k)$ were determined from the statistical characteristics of the investigated segments of the speech signal on the basis of generalizing the method of linear prediction from the vector case [1]. Minimization of the variance of the prediction error

$$E_j^2(k) = M\left\{\left(x(k) - F_j(k,k-1)\, x(k-1)\right)\left(x(k) - F_j(k,k-1)\, x(k-1)\right)^T\right\},$$

leads to the condition of orthogonality

$$M\left\{\left(x(k) - F_j(k,k-1) x(k-1)\right) x^T(1)\right\} = G_j(k)\, M\,[W(k) W^T(k)]\, G_j^T(k)\, \delta_{k,l},$$

REFERENCES

1. J. McHowell, "Linear prediction. Review," Proc. IEEE [in Russian], vol. 63, no. 4, pp. 20-45, 1975.

2. M. A. Mironov and M. S. Yarlykov, "Optimal discrete algorithms of functional diagnosis of the technical state of dynamic systems," Avtomatika i Telemekhanika, no. 10, pp. 144-151, 1985.

3. M. A. Mironov, "Optimal discrete processing of signal dependent simultaneously on several samples of the stochastic processes being estimated," Radiotekhnika i Elektronika, vol. 37, no. 1, pp. 107-116, 1992.

4. S. Ya. Zhuk, "Synthesis of digital detector-meters of mixed Markov processes," Izv. VUZ. Radioelektronika, no. 11, pp. 31-37, 1989.

5 May 1992