Radioelectronics and Communications Systems
Vol. 47, No. 3, pp. 46–48, 2004

Izvestiya VUZ. Radioelektronika
Vol. 47, No. 3, pp. 66–70, 2004
UDC 623.624.347:681.32

A TEXT-INDEPENDENT SYSTEM OF USER VOICE IDENTIFICATION (AUTHENTIFICATION)

L. K. Babenko, O. B. Makarevitch, V. M. Fyodorov, and P. Yu. Yurkov

Taganrog State Radio-Engineering University, Russia

The paper considers peculiarities of authentication (identification) of a user by his speech signal based on combination of methods of fast digital processing of signals and neural networks.

The main elements of contemporary systems of access control are authentication and identification of a user, procedures of access control, methods of access to protected data and, eventually, the methods of reporting and auditing. All these elements are well studied and formalized. As shown by investigation, the largest number of violations of the safety system occurs at the stage of identification, i.e., recognition of the user's personality by the authenticator produced, and at the stage of user authentication (verification), i.e., checking whether the presented personal data correspond to any registered users. Today for authentication (identification) purposes the systems of names and passwords or electronic cards (TouchMemory, DallasLock, etc.) are used.

These systems are assumed to have a rather high level of protection. Biometric methods of identification (authentication), for example, by finger or palm prints, or by human retina, are no less reliable. However, the enumerated methods cannot be made reticent and changed in case of necessity. As a result, they are highly vulnerable to forgery. The authentication (identification) of the user by his voice is free from these disadvantages, but realization of this technique presents considerable difficulties: the properties of a speech signal are such that the use of only classical methods of digital signal processing does not permit authentication (identification) with high accuracy and speed. In the general case, identification by voice is a rather sophisticated problem, and cannot be resolved without application of elements of artificial intelligence.

The purpose of this work is to suggest another approach to resolving this task based on combination of the methods of high-speed digital processing of signals and of neural networks.

The problem of user authentication is to determine whether the presented name and voice sample correspond to those of the user registered in the system. However, direct comparison is impossible for a number of reasons, particularly, because of random nature of distortion during the recording, and also due to irregularities of the speech signals in amplitude and time. Thus, in text-independent authentication of a person by his voice we have to use a set of some stable parameters picked out of the speech signal followed by their processing using artificial neural networks.

It is known that the information inherent in a particular speaker and contained in the speech signal is mainly concentrated in the middle and upper parts of the spectrum. Remembering that speech is a multiplicative signal, in which one cofactor represents a rapidly varying, while the second — slowly varying signal, their separation is performed by nonlinear methods, as a rule, by homomorphic processing of speech [1]. Usually the calculation of cepstrum is performed with the inverse Fourier transformation of the logarithm of signal power spectrum [2]. However, in [2] the cepstral coefficients (CC) were calculated recursively by the linear transform coefficients. If the linear prediction filter is stable (its stability is guaranteed when using the autocorrelation method), the logarithm of the reverse filter can be expressed as an energy series in z^{-1} [3]:

Radioelectronics and Communications Systems
Vol. 47, No. 3, 2004

REFERENCES

1. L. R. Rabiner and R. V. Shafer, Digital Processing of Speech Signals [in Russian], Radio i Svyaz', Moscow, 1981.
2. G. Markel and A. Grey, Linear Prediction of Speech [Russian translation ed. by Yu. N. Prokhorov and V. S. Zvezdin], Svyaz', Moscow, 1980.
3. S. L. Marple, Jr., Digital Spectral Analysis with Applications, Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1987.
4. B. S. Atal and S. L. Hanauer, Journal of the Acoustic Society of America, Vol. 50, No. 2, pp. 637–655, 1971.
5. B. S. Atal, Journal of the Acoustic Society of America, Vol. 55, No. 6, pp. 1304–1311, 1974.
6. A. V. Ramishwili, Automatic Identification of the Speaker [in Russian], Mir, Moscow, 1989.

3 November 2003