M. M. Bykov, Sc. (Eng.), Prof.; V. V. Kovtun, Sc. (Eng.), Assist. Prof.; N. F. Savinova

EVALUATION OF NOISE INFLUENCE ON THE RELIABILI INFORMATION-MEASURING VOICE RECOGNITION SYSTEM OPERATION

The paper considers the influence of noise in the speech signal on the reliability of the information-measuring system of voice recognition operation. Analytical expressions for the evaluation of recognition errors are obtained , calculation results of the probability of error classification of two classes of voices at different levels of noise in the signal are given .

Key words: *voice recognition, information-measuring system, reliability, noise, mathematical model, transmission channel, distribution hyperplane, normal distribution law, rule of minimum distance.*

The transmission channel through which the distribution of speech signals occurs, will be used for voice recognition is influenced by all kinds of noise [1], main among them are the noise of equipment and the environment. The given influence leads to decrease of the reliability of the information-measuring system for voice recognition operation . Therefore, evaluation of this impact, taking into account the structural features of information-measuring system of this type is an urgent task.

Using the linear model [2, 3, 4], the speech signal can be regarded as quasi-determined process for the same classes – images at time interval of $\tau = 10 \div 30$ milliseconds. In the presence of additive noise speech signal will have the following form:

$$y^*(t,\tau) = y(t,\tau) + \xi(t,\tau). \tag{1}$$

In most cases the noise $\xi(t,\tau)$ has zero average value and is non-correlated with the speech signal $y(t,\tau)$. Consequently, the vector of the signal can be represented as:

$$X = X_y + X_\xi, \tag{2}$$

where $X_y = \left(x_{y_1}, x_{y_2}, \ldots, x_{y_n}\right)^T$, $X_\xi = \left(x_{\xi_1}, x_{\xi_2}, \ldots, x_{\xi_n}\right)^T$ – are the vectors of speech signal description and noise respectively ;

T – is symbol of transposition ;

$x_i = x_{y_i}^2 + x_{\zeta_i}^2$ – is spectral power at the i^{th} frequency band.

For the leveling of the influence of speech signal volume on the results of recognition , we use normalized description vector

$$\|X\| = \sum_{i=1}^{n} x_i = \sigma_c^2, \tag{3}$$

where σ_c^2 - is the dispersion at the input of the device of spectral analysis of the signal.

As a result of vector (2) normalization , respectively, from (3) we obtain:

$$\widetilde{X} = \frac{r^2}{r^2+1}\widetilde{X}_y + \frac{1}{r^2+1}\widetilde{X}_\xi,$$

where $\widetilde{X}, \widetilde{X}_y, \widetilde{X}_\xi$ are normalized vectors X, X_y, X_ξ respectively,

$r = \sqrt{\dfrac{\sigma_y^2}{\sigma_\xi^2}}$ –is "signal / noise” ratio at the input of recognition system.

Taking into account that in practice, $r^2 >> 1$ the equation will have the form:

$$\widetilde{X} = \widetilde{X}_y + \frac{1}{r^2}\widetilde{X}_\xi. \qquad (4)$$

Operation reliability of information-measuring system classifier for voice recognition will be defined define on the example of two classes of speakers Ω_1 and Ω_2 then it will not be difficult to generalize for the case of larger number of classes. A typical classifier, which is used in information-measuring system for voice recognition, is a classifier " by minimum of distance" [5]. Hence, for this type of classifier classification rule has the form:

$$d_e\left(\widetilde{X}, \mu_i\right) = \min d_e\left(\widetilde{X}, \mu_j\right) \Rightarrow X \in \Omega_i, \qquad i, j = 1,2, \qquad (5)$$

where $\mu_i = E^0\left\{\widetilde{X} \mid \widetilde{X} \in \Omega_i\right\}$ is the average value of vectors $\widetilde{X}$, which belong to class Ω_i (Ω_i is the reference of the class);

$d_e\left(\widetilde{X}, \mu_i\right) = \left|\left(\widetilde{X} - \mu_i\right)^T\left(\widetilde{X} - \mu_i\right)\right|$ is Euclidean distance from vector $\widetilde{X}$ to vector μ_i.

In terms of hyperplane that separates classes of speakers Ω_1 and Ω_2, and the rule (5) for signal without noise will have the form:

$$H_{12}\left(\widetilde{X}\right) = \left(\widetilde{X}_y - \mu_1\right)^T\left(\widetilde{X}_y - \mu_1\right) - \left(\widetilde{X}_y - \mu_2\right)^T\left(\widetilde{X}_y - \mu_2\right) = 0, \qquad (6)$$

or

$$H_{12}\left(\widetilde{X}\right) = 2\widetilde{X}_y^T\left(\mu_2 - \mu_1\right) + \mu_1^T\mu_1 - \mu_2^T\mu_2 = 0. \qquad (7)$$

Having substituted (4) into (7), we will form the equation of the hyperplane in the presence of noise, assuming that training of the classifier was performed on the signal without noise

$$H_{12}^r\left(\widetilde{X}\right) = 2\widetilde{X}_y^T\left(\mu_2 - \mu_1\right) + \frac{2}{r^2}\widetilde{X}_\xi^T\left(\mu_2 - \mu_1\right) + \mu_1^T\mu_1 - \mu_2^T\mu_2 = 0, \qquad (8)$$

or

$$H_{12}^r\left(\widetilde{X}\right) = H_{12}\left(\widetilde{X}\right) + \frac{2}{r^2}\widetilde{X}_\xi^T\left(\mu_2 - \mu_1\right) = 0.$$

From equation (8) it is seen that the effect of noise leads to the shift of the hyperplane to one of the classes Ω_1 or Ω_2, depending on the sign of the vector $\left(\mu_2 - \mu_1\right)$. Since quasidetermined speech signal power of instanteneous spectrum in the i^{th} frequency band is a random function [6], the spectral description vector $\widetilde{X}$ is a random vector, which is characterized by a multidimensional normal distribution. Then (7), (8) are defined by the density of one-dimensional normal distribution. The average value for the decision function (7) will be:

$$E^0\left\{H_{12}\left(\widetilde{X}\right)\right\} = 2E^0\left\{\widetilde{X}_y\right\}\left(\mu_2 - \mu_1\right)^T + \mu_1^T\mu_1 - \mu_2^T\mu_2. \qquad (9)$$

And for the decision function (8), assuming $\frac{2}{r^2}E^0\left\{\widetilde{X}_\xi^T\right\} \approx \frac{2}{r^2}\widetilde{X}_\xi^T$,

$$E^0\left\{H_{12}\left(\widetilde{X}\right)\right\} = 2E^0\left\{\widetilde{X}_y\right\}\left(\mu_2 - \mu_1\right)^T + \frac{2}{r^2}\widetilde{X}_\xi^T\left(\mu_2 - \mu_1\right) + \mu_1^T\mu_1 - \mu_2^T\mu_2, \qquad (10)$$

Dispersion of the decision function (7) will be:

$$\sigma_н^2 = E^0\left\{\left[2\widetilde{X}_y\left(\mu_2 - \mu_1\right)^T + \mu_1^T\mu_1 - \mu_2^T\mu_2\right] - \left[2E^0\left\{\widetilde{X}_y\right\}\left(\mu_2 - \mu_1\right)^T + \mu_1^T\mu_1 - \mu_2^T\mu_2\right]\right\}^2, \qquad (11)$$

After simplification we obtain

$$\sigma_{н}^{2} = 2(\mu_2 - \mu_1)^T \sum_{i}^{0} (\mu_2 - \mu_1), \qquad i = 1,2, \tag{12}$$

where $\sum\limits_{i}^{0}$ – is covariance matrix of the i^{th} class.

For simplification of the expression we assume $\sum\limits_{1}^{0} = \sum\limits_{2}^{0} = \sum\limits^{0}$.

Having substituted (8) into (11), we obtain the equation of the decision function $H_{12}^{r}(\widetilde{X})$, analogous to equation (12).

Consequently, when $\widetilde{X} \in \Omega_1$ the decision functions $H_{12}(\widetilde{X})$ and $H_{12}^{r}(\widetilde{X})$ are calculated by normal laws $N_1(m_1^0, \sigma_н)$ and $N_1^r\left(m_1^0 + \frac{2}{r^2}\widetilde{X}_{\xi}^{T}(\mu_2 - \mu_1), \sigma_н\right)$, respectively. The average value of m_1^0 we obtained substituting in (11) the value $E^0\{\widetilde{X}_y\} = \mu_1$:

$$m_1^0 = \mu_2^T (2\mu_1 - \mu_2) - \mu_1^T \mu_1. \tag{13}$$

Similarly, if $\widetilde{X} \in \Omega_2$, decision functions $H_{12}(\widetilde{X})$ and $H_{12}^{r}(\widetilde{X})$ are calculated by normal laws $N_2(m_2^0, \sigma_н)$ and $N_2^r\left(m_2^0 + \frac{2}{r^2}\widetilde{X}_{\xi}^{T}(\mu_2 - \mu_1), \sigma_н\right)$, respectively. The average value of m_2^0 we obtained substituting in (9) the value $E^0\{\widetilde{X}_y\} = \mu_2$:

$$m_2^0 = \mu_1^T (\mu_1 - 2\mu_2) - \mu_2^T \mu_2. \tag{14}$$

Thus, the probability of errors of the first and second order for decision function (17) in the absence of noise are calculated by the equation

$$P(e) = P(\Omega_1) P\left(H_{12}(\widetilde{X}) < \theta \,|_{\widetilde{X} \in \Omega_1}\right) + P(\Omega_2) P\left(H_{12}(\widetilde{X}) > \theta \,|_{\widetilde{X} \in \Omega_2}\right), \tag{15}$$

where θ – is the threshold of recognition.

Taking into account (13) and (14), we obtain::

$$P\left(H_{12}(\widetilde{X}) < \theta \,|_{\widetilde{X} \in \Omega_1}\right) = \int_{-\infty}^{\theta} \frac{1}{\sigma_н \sqrt{2\pi}} e^{\frac{-\left[H_{12}(\widetilde{X}) - m_1^0\right]}{2\sigma_н^2}} dH_{12}(\widetilde{X}), \tag{16}$$

$$P\left(H_{12}(\widetilde{X}) > \theta \,|_{\widetilde{X} \in \Omega_2}\right) = \int_{\theta}^{\infty} \frac{1}{\sigma_н \sqrt{2\pi}} e^{\frac{-\left[H_{12}(\widetilde{X}) - m_2^0\right]}{2\sigma_н^2}} dH_{12}(\widetilde{X}). \tag{17}$$

Substituting (16), (17) into (15), we obtain:

$$P(e) = P(\Omega_1) \Phi^{*}\left(\frac{\theta - m_1^0}{\sigma_н}\right) + P(\Omega_2) \Phi^{*}\left(\frac{m_2^0 - \theta}{\sigma_н}\right), \tag{18}$$

where $\Phi^{*}(a)$ –is Laplace function.

Choosing binary function of losses (0 - correct recognition, 1 - error), the threshold value θ will be determined by the ratio:

$$\frac{P\left(H_{12}\left(\widetilde{X}\right)=\theta\,|_{\widetilde{X}\in\Omega_1}\right)}{P\left(H_{12}\left(\widetilde{X}\right)=\theta\,|_{\widetilde{X}\in\Omega_2}\right)}=\frac{P(\Omega_1)}{P(\Omega_2)}=\theta_0. \qquad (19)$$

Substituting (15) and (16) into (19), we obtain :

$$\theta=\frac{m_1^0+m_2^0}{2}+\frac{\sigma_н^2\ln\theta_0}{m_1^0-m_2^0}. \qquad (20)$$

Replacing in (20) values, m_1^0, m_2^0 and $\sigma_н^2$ by their meanings, we obtain the threshold value of θ

$$\theta=\sum^{0}\ln\theta_0. \qquad (21)$$

In case of noise threshold value θ is written as:

$$\theta_\xi=\frac{4}{r^2}X_\xi^T\left(\mu_2-\mu_1\right)+\sum^{0}\ln\theta_0=\Delta_\xi+\theta. \qquad (22)$$

Taking into account (22), the probability of errors of the first and second order for decision function (18), in case of noise present, will be:

$$P\left(e_\xi\right)=P(\Omega_1)\Phi^*\left(\frac{\theta+\Delta_\xi-m_1^0}{\sigma_н}\right)+P(\Omega_2)\Phi^*\left(\frac{m_2^0-\Delta_\xi-\theta}{\sigma_н}\right). \qquad (23)$$

Equation (23) determines the dependence of recognition error on the presence of noise in the speech signal. The advantage of the proposed method of the account of the impact of noise available in speech signal on the operation reliability of the information-measuring system for voice recognition is the fact that there is no need to calculate the integral of the probability distribution density of information characteristics features in n --dimensional space.

Using formulas (18) and (23), the study of the dependence of voice recognition error on noise available in the speech signal is carried out. The results of research are presented in Fig. 1.

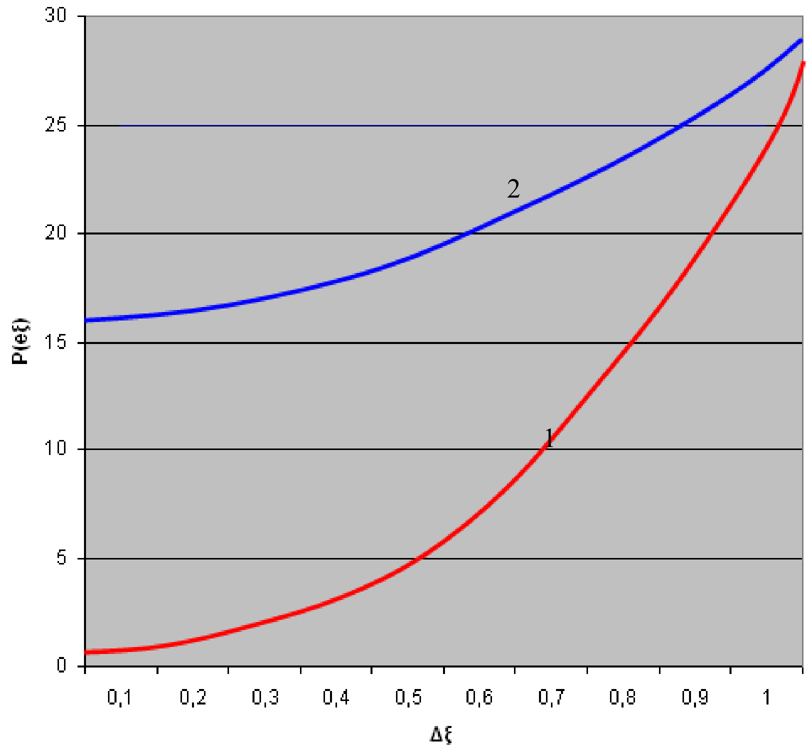

Fig. 1. Dependence of recognition errors on the magnitude of noise available in the speech signal: 1: 2:

$m_1^0 = 1,\ \sigma_и = 1;$ 2: $m_1^0 = 1,\ \sigma_и = 0{,}5$

CONCLUSIONS

As a result of research mathematical model of the influence of noise on the reliability of voice recognition, which eliminates the need for integration of erroneous decisions area into a multidimensional characteristic space is elaborated. Analytical expositions, carried out and experimental results showed:

– reduction of noise influence on the operation reliability of the information-measuring system for voice recognition applying the method of their filtration is efficient for the information characteristics with high separation of speakers classes, that is, if $\sigma_и \le {m_1^0}/{2}$;

– filtration of noise with large amplitude is more efficient while voice recognition, than filtration of errors with small amplitude.

REFERENCES

1. Давенпорт В. Б., Рут В. Л. Введение в теорию случайных сигналов и шумов. – М.: ИЛ, 1970. – 498 с.
2. Рамишвили Г.С. Автоматическое опознавание говорящего по голосу. – М.: Радио и связь, 1981. – 224 с.
3. Фант Г. Акустическая теория речеобразования. – М.: Наука, 1964. – 284 с.
4. Рабинер Л. Р. Шафер Р. В. Цифровая обработка речевых сигналов: Пер. с англ. – М.: Радио и связь, 1981. – 496 с.
5. Ротштейн А. П. Интеллектуальные технологии идентификации: нечеткие множества, генетические алгоритмы, нейронные сети. – Винница: УНИВЕРСУМ-Вінниця, 1999. – 320 с.
6. Харкевич А. А. Спектры и анализ. – М.: Физматиз, 1962. – 320 с.

Bykov Mykolay – Cand.Sc.(Eng), Professor, Department of Computer Control Systems.

Kovtun Vyacheslav – Cand.Sc. Assist. Professor, Department of Computer Control Systems.

Savinova Natalia – Post-graduate, Department of Computer Control Systems.
Vinnytsia National Technical University.