TECHNOLOGY IS GETTING RID OF THE NOISE IN SPEECH PERCEPTION

I. U. Nazarov, Senior lecturer, Department of “Information technologies” Tashkent state pedagogical university named after Nizami

M. M. Payazov, Doctorate of TUIT (Tashkent University of Information Technologies)

M. Z. Tadjibayeva, lecturer of the Department “Methodology of Teaching Computer science”, Tashkent state pedagogical university named after Nizami

ABSTRACT

The article presents the technology of processing human speech by processing information on an extended frequency range, which gives the potential to use it in different devices transmitted through communication channels, IP-telephony, telecommunications and in the field of medicine, namely - in audiology and sign language.

Keywords: Speech signals, phoneme, filtering, noise, acoustics, frequency range.

INTRODUCTION, LITERATURE REVIEW AND DISCUSSION

The development of information and communication technologies opens up many new opportunities for the processing of speech signals. When a person perceives speech in the field of telecommunications technology or in the field of medicine as audiology and audiology with the help of special devices, something often happens in which the person does not hear well or the device works incorrectly and makes strange sounds. And in fact, this can happen because of the error and the occurrence of different noisy parameters on these devices.

Voice signals used on different devices are always more or less noisy. In order for human speech in these devices to be clear and clear, signal-processing technology is needed. To date, in the field of speech recognition and identification, the development of noise reduction methods or noise filtering is an important area of research. The term "noise" refers to unwanted electrical signals that are always present in electrical systems[1].

The presence of noise superimposed on the signal masks it, this limits the receiver's ability to make accurate decisions about the meaning of the symbols, and therefore limits the speed of information transmission. Noise can be natural and artificial. Natural noises come from the atmosphere, the sun and other galactic sources. Artificial noise is spark ignition noise, switching pulse noise, and noise from other related sources of electromagnetic radiation.

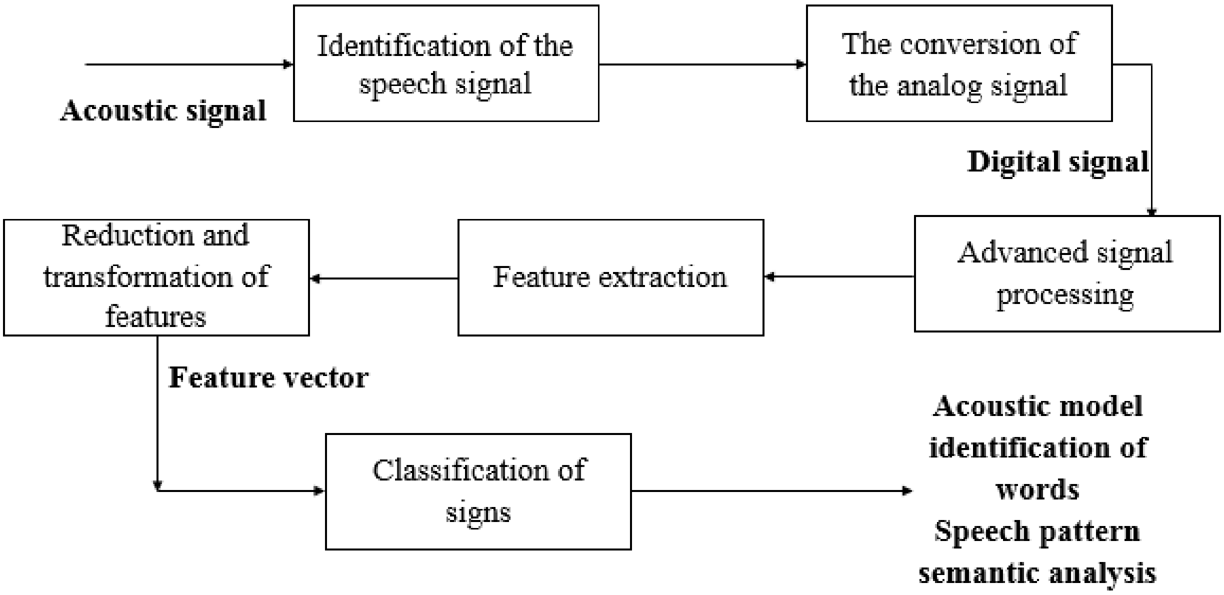
In speech technology, there are many different noise reduction algorithms. The most common noise model for the algorithms under consideration is additive white Gaussian noise *(abgs)*. Also, there are different methods to determine speech intelligibility[2]:

---

[1] S. Bernard "Digital communication" theoretical foundations and practical application, Moscow, St. Petersburg, Kiev-2003., pp. 58-60.
[2] S. A. Novoselov, A. I. Topnikov, A. I. Savvatin, A. L. Priorov "noise Suppression in speech signals based on the nonlocal averaging method" - Digital signal processing No. 1/2012, page-23.

- methods based on the use of autoregressive model of speech signal;
- methods based on speech signal processing using hidden Markov models;
- methods based on artificial neural networks;
- methods based on estimation of noise parameters of mean square error minimization and threshold processing in the transform ant region;
- methods for adaptive interference cancellation;
- methods based on the use of mathematical models of speech signals in the time domain;
- methods based on the use of spectral characteristics of noise and others.

In the new technology, getting rid of noise in speech perception on the example of Uzbek speech provides for the recognition of the speech signal by extracting the necessary features for identification and comparing them with the corresponding database of images *(Rice. 1.).*



Rice. 1. Recognition algorithm of the speech signal

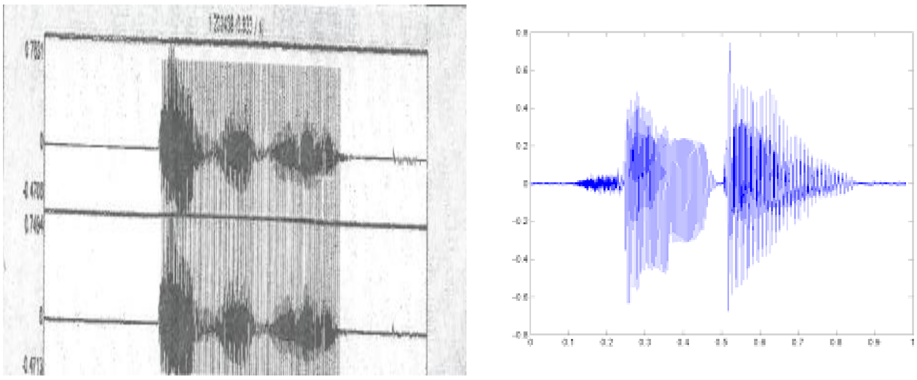
The acoustic signal enters the speech signal identification unit, then the analog signal in the conversion unit is converted into a digital signal, which enters the preprocessing unit, then the extraction of features is carried out, after which the reduction and transformation of features is carried out, the process is converted to feature vectors and feature classification, resulting in an acoustic model, word identification, speech model and semantic analysis.

In this case, the recognition process is carried out sequentially from the fixed set of signs of the signal through the tested hypothesis of the assumed information to a sufficiently reliably recognized speech signal.

The perception of speech with the help of certain communication channels or other special devices such as hearing AIDS will often inevitably lead to distortion or complete failure of information. To obtain high results, studies were conducted with the help of different speech audio recordings by speakers of different gender and age in the Uzbek language. Similar recordings in other languages were used for control. It was also compared with similar signals with a reduced frequency range.

In the process of speech recognition and identification, the main frequency of speech plays an important role. In the speech signal, there was a deviation of the male voice in the frequency range from 60 to 250 Hz.

In the existing methods of analysis of speech signals were guided by the range from 50 Hz to 4 KHz, considering that beyond this frequency range is located only not essential for speech recognition information. All studies were extended to the entire range on the example of human hearing aid perception with a frequency range from 20 Hz to 20 KHz in normal speech with a sound strength of 80 dB by Uzbek speakers *(Rice.2.)*.



Rice. 2. Analysis of Uzbek speech on an extended frequency range.

As a result, it was determined that the basic tone of Uzbek speech is a lower basic tone compared to other languages. Recognition of components of Uzbek speech audio signals in an extended frequency range, including audio signal processing, conversion to a digital signal, characterized in that the digital signal enters the pre-processing unit, then carry out the extraction of signs, then carry out the reduction and transformation of signs, the process is converted to the vector of signs, carry out the classification of signs.

The objective of the present study is to improve the quality and intelligibility of speech perception against the background of noise on various radio, telephony, telecommunications, television channels, including modern digital hearing AIDS.

This study is based on the analysis of multicomponent speech signals in the extended frequency range. The advantage of this study is that it allows to improve the quality of perception of any speech signals with an extended frequency range.

To obtain high results, studies were conducted with the help of different speech audio recordings by speakers of different gender and age in the Uzbek language. Similar records in romance and Slavic languages were used for control. It was also compared with similar signals with a reduced frequency range.

Generally accepted law in the recognition and identification of speech plays an important role basic frequency of speech. In the speech signal, there was a deviation of the male voice in the frequency range from 60 to 250 Hz. With existing methods of analysis of speech signals, be guided by the range from 50 Hz to 4 KHz and considering that beyond this frequency range is located exclusively not essential for speech recognition information. When determining words and phrases, it was revealed that the frequency range of the first formant was located below the frequency of the main background. The decrease in the basic tone in the Uzbek language was

ISSN 2056-5852

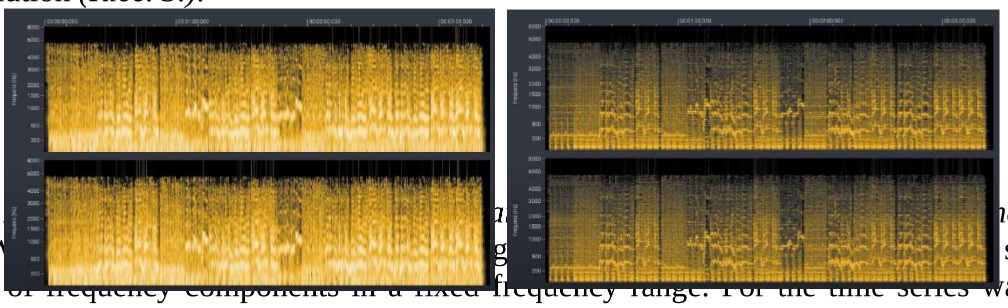
also accompanied by similar decreases in frequency ranges for formants. If the pronunciation of the vocal "a" in the romance languages, the first formant averaged 1000 Hz, then in the Uzbek speech it fluctuated within 700-800 Hz depending on the degree of stress. And for the second formant, there was no significant difference between speakers of different language backgrounds. However, such discrepancies did not occur in all vocals, for example, in the vocals "o" was fixed full compliance for all the studied language groups. The specific sounds "u" and" I "in parameters corresponded to the umlaut of the German speech "ü" " ä "while the umlaut" ö " due to its phonetic features has no analogues.

An algorithm has been developed to improve the quality of speech signals. In practice, high frequency ranges are not paid attention to in principle, as it is believed that they do not play any significant role in speech recognition and identification. High frequency bands are also not included in the encoding of digital voice signals. Thus, to ensure the compactness of the speech audio file, they will usually be overboard digital information. As a result of research on hissing sound, it was revealed that the frequency structure and combinations of format changes were fundamentally the same for both male and female speakers.

Most computer programs are designed to analyze and interpret the time and frequency content of acoustic signals. Typical types of analysis in these computer programs are spectral analysis, wavelet analysis, spectral analysis, and spectral histogram.

Spectrum analysis represents each acoustic signal as the sum of frequency components in a fixed frequency range. The time series transformation will be implemented using the fast Fourier Transform *(FFT)* mathematical tool. The spectrogram will be created as spectra of slices of the acoustic signal at regular intervals and displays the signal frequencies at the time of the signal duration.

Graphical display spectrograms show the frequent ranges where most of the sound energy is concentrated. The frequency and time resolution can be changed by determining the length of each slice. Short time slices give better resolution for the duration of time, but blur the resolution of frequencies. Longer time intervals give better frequency resolution with blurred time resolution *(Rice. 3.)*.

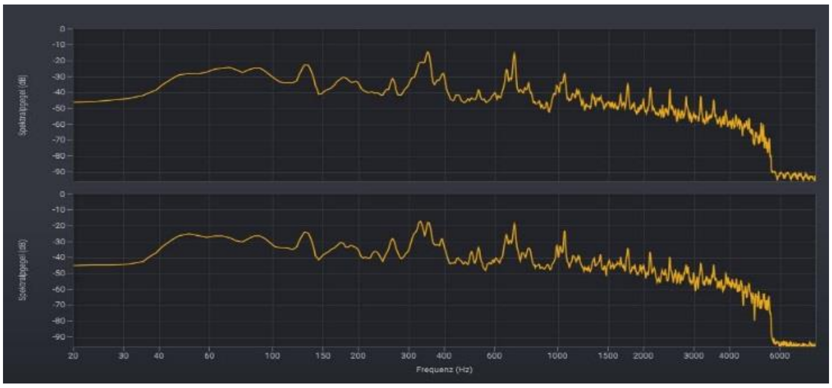


[illegible] a[illegible] nd.

A W[illegible] g[illegible] signal as a sum [illegible] wavelet transform, of course, the wavelet transform is used.

The most distinctive feature of the spectrogram is the dependence of the cut length on the frequency. As a result, the higher frequencies of the signal get a better temporal resolution, but with a blurred resolution of the lower frequencies.

For acoustic signal analysis, the loudness timeline information is not representative of any deductions. More important, however, is the relationship between sound intensity and frequencies. Thus, spectral analysis shows how the intensity depends on the frequencies inside the signal. As a result of the analysis, the frequency distribution range can be determined *(Rice. 4.)*.



Rice. 4. Spectral analysis of the compressed sound (vocal with instrumental accompaniment).

Taking into account the above, it can be clearly established that for adequate recognition and identification of speech, it is necessary to study them in an expanded frequency range and implement them in practice.

At the same time, the study shows that taking into account the behavior of the factors of the speech signal during its sounding, each background was characterized by certain numerical values, which, taking into account some deviations, formed a kind of signature.

This circumstance emphasizes the principal possibility to use for identification and recognition of speech signals methods of recognition of samples that give more correct results than currently used probabilistic methods.

Literature

1. Bernard S. "Digital communication" theoretical foundations and practical application, Moscow, St. Petersburg, Kiev-2003., pp. 58-60.
2. Novoselov S.A., Topnikov A.I., Savvatin A.I., Priorov A.L. "noise Suppression in speech signals based on the nonlocal averaging method" - Digital signal processing No. 1/2012, page-23.
3. Khasanov A.A. Didactic Foundations of Interdisciplinary Connections at Subject Teaching // Eastern European Scientific Journal. Germany -2018. No. 6. P. 127-130.
4. Nazarov U., Payazov M. M. The study of the degree of accuracy of speech signals in the process of local and global inter-network exchange of information. Pedagogy, 4/2019. Tashkent 2019.
5. Hasanov A.A., Gatiyatulina R.M. Interdisciplinary Communication as a Didactic Condition of Increasing the Efficiency of Educational Process // Eastern European Scientific Journal. Germany -2016. No. 5. P. 107-111.
6. Mamarajabov O.E. Benefits of Using Information Technology in the Education System // Vocational Education. Tashkent, 2019. No.1. P. 55-59
7. Urokova Sh.B. Tuhtashev U.F. Trends of electronic education development // European Journal of Research and Reflection in Educational Sciences Vol. 7. No.12, 2019. P. 485-488.
8. Bakiyeva F.R, Mirzahmedova N.D. Efficiency of online training // International Scientific Journal Theoretical & Applied Science. Vol. 79. Issue:11. Year: 2019. P. 56-58.