International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Definition of Speech Intelligibility of the Uzbek Language

Narzillo Mamatov, Mirabbos Payazov, Nilufar Niyozmatova, Yusuf Yuldoshev , Naibakhon Mamadalieva

***Abstract*: This article proposes an algorithm for automating the process of personality recognition based on voice, provides an analysis of existing methods used to solve the problem that needs to be solved. A method was implemented based on the Gaussian mixture model, which distinguishes a person's voice with high accuracy. The components of this model allow you to simulate sound characteristics that are unique to each person. The results of the proposed algorithm and the use of voice recognition based on the results of the proposed algorithm are presented.**

***Keywords*: Speech signals, phoneme, filtering, noises, acoustics, frequency range.**

I. INTRODUCTION

The speech signals used on different devices are always noisy to one degree or another. For human speech in these devices to be clear and clear, signal processing technology is needed. Today, in the field of speech recognition and identification, the development of noise reduction methods or noise filtering is an actual area of research.

The presence of noise superimposed on the signal obscures or masks the signal, this limits the receiver's ability to make accurate decisions about the meaning of the symbols, and therefore limits the transmission rate of information. Noises are natural and artificial. Natural noise comes from the atmosphere, the sun, and other galactic sources. Artificial noise is spark ignition noise, switching impulse noise and noise from other related sources of electromagnetic radiation.

Currently, there are many different noise reduction algorithms for speech technologies. A common noise model for such algorithms is additive white Gaussian noise. Classification of methods for improving the quality and intelligibility of speech is carried out on the basis of the basic concept and the idea underlying the method.

Revised Manuscript Received on January 30, 2020.
* Correspondence Author

Mamatov Narzillo*, Tashkent University Information Technologies named after Al-Kharezmi, Tashkent, Uzbekistan.

Payazov Mirabbos, Tashkent University Information Technologies named after Al-Kharezmi, Tashkent, Uzbekistan.

Niyozmatova Nilufar, Tashkent University Information Technologies named after Al-Kharezmi, Tashkent, Uzbekistan.

Yuldoshev Yusuf, Tashkent University Information Technologies **named** after Al-Kharezmi, Tashkent, Uzbekistan.

Naibakhon Mamadalieva, Uzbekistan National University of Uzbekistan named after Mirzo Ulugbek, Tashkent, Uzbekistan.

The following groups of methods are currently widespread [1-4]:

- based on the use of an autoregressive model of a speech signal;
- based on the processing of a speech signal using the apparatus of hidden Markov models;
- based on artificial neural networks;
- based on the estimation of noise parameters, minimizing the standard error and threshold processing in the field of transformants.
- adaptive interference compensation;
- based on the use of mathematical models of speech signals in the time domain;
- based on the use of spectral characteristics of noise and others.

II. STATEMENT AND SOLUTION OF THE PROBLEM

A new technology for noise purification in speech perception using the example of Uzbek speech provides for the recognition of a speech signal by extraction of signs necessary for identification and their comparison with the corresponding database of images (Fig-1).

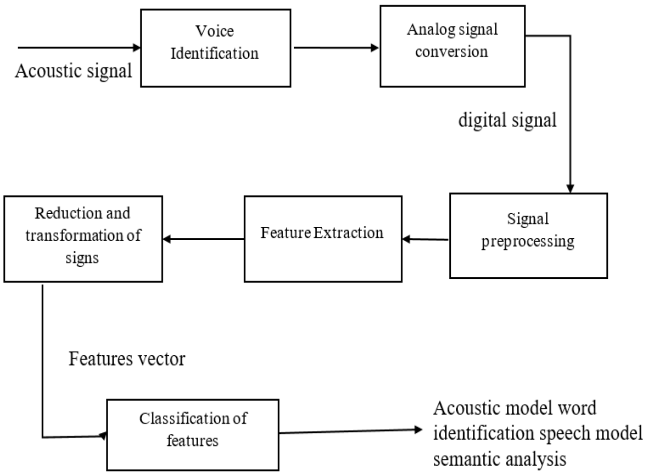

Fig-1. Speech recognition algorithm

The acoustic signal is transmitted to the speech signal identification unit, then the analog signal in the conversion unit is converted to a digital signal, which is fed to the pre-processing unit, then the features are extracted and then the features are reduced and the transformation is converted to the feature vectors and the classification of the signs results in an acoustic model, word identification, speech model and semantic analysis.

In this case, the recognition process is carried out sequentially from a fixed set of signal attributes through the tested hypothesis of the alleged information to a fairly reliably recognizable speech signal (Fig-2).

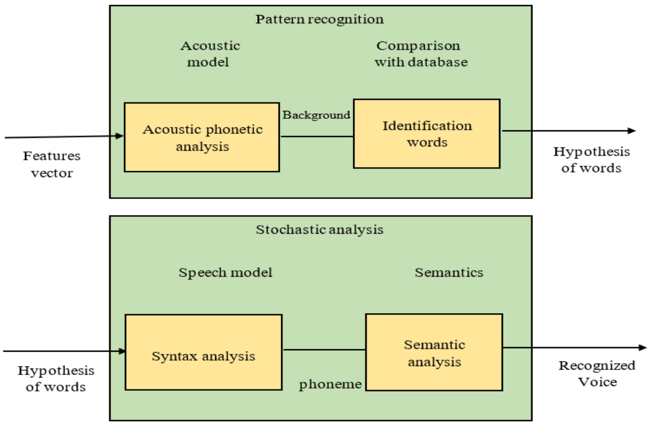

Fig-2. The scheme of analysis and recognition of speech signals

The perception of speech using certain communication channels or other special devices, such as hearing aids, often leads without fail to distortion or to a complete failure of information.

To obtain high results, studies were conducted using various voice audio recordings by speakers of different sex and age in the Uzbek language. For control similar records in other languages were involved. A comparison was also made 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, a deviation of the male voice was observed in the frequency range from 60 to 250 Hz.

With existing methods for analyzing speech signals, they were guided by a range from 50 Hz to 4 KHz, considering that outside this frequency range there was exclusively information that was insignificant for speech recognition. All studies were extended to the entire range by the example of perception by a human hearing aid with a frequency range from 20 Hz to 20 KHz in normal speech with a sound power of 80 db by Uzbek speakers.

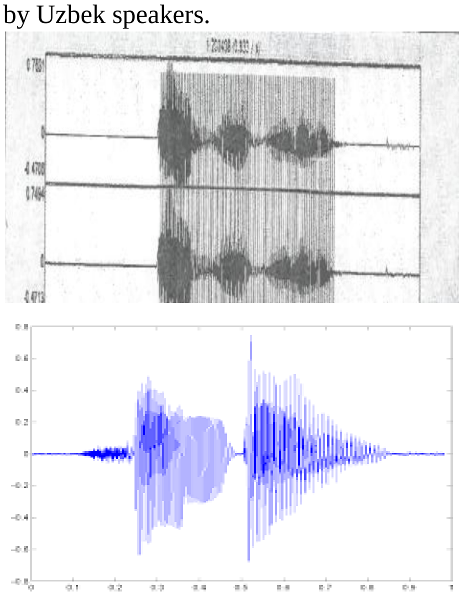

Fig-3. Analysis of Uzbek speech in an extended frequency range.

III. RESULT AND DISCUSSION

The results of the experiments showed that the main tone of the Uzbek language in comparison with the speeches of other languages is a lower main tone.

Recognizing the components of the Uzbek speech audio signals in the extended frequency range includes processing the audio signal, as well as converting it to a digital signal, characterized in that the digital signal is fed to the pre-processing unit, then the signs are extracted, then the signs are reduced and transformed, the process converts signs into a vector, at the end the classification of signs is carried out.

The objective of this study is to provide improved quality and intelligibility of speech perception against the background of noise on different radio channels, telephony, telecommunications, television, including modern digital hearing aids.

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

To obtain high results, studies were conducted using various voice audio recordings by speakers of different sex and age in the Uzbek language. For control, similar records were used in the Romance and Slavic languages. A comparison was also made with similar signals with a reduced frequency range.

The generally accepted law of speech recognition and identification plays a major role in the main frequency of speech. In the speech signal, a deviation of the male voice was observed in the frequency range from 60 to 250 Hz. With existing methods for analyzing speech signals, be guided by the range from 50 Hz to 4 KHz and believing that outside this frequency range there is exclusively information insignificant for speech recognition. When defining words and phrases, such a circumstance 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 also accompanied by similar decreases in the frequency ranges according to formants. If, when pronouncing vocal “a” in Romance languages, the first formant averaged 1000 Hz. In Uzbek speech, he ranged between 700-800 depending on the degree of stress. And according to the second formant, there were no significant differences between speakers of different language affiliations. However, such discrepancies did not occur for all vocals, for example, for vocals “o”, full correspondence was recorded for all studied language groups. The specific sounds “y” and “I” in terms of parameters corresponded to the umlaut of the German speech “ü” “ä” while the umlaut of “ö”, due to its phonetic features, has no analogues.

To improve the quality of speech signals compiled an algorithm.

Retrieval Number: B6370129219/2020©BEIESP
DOI: 10.35940/ijitee.B6370.019320
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication (BEIESP)

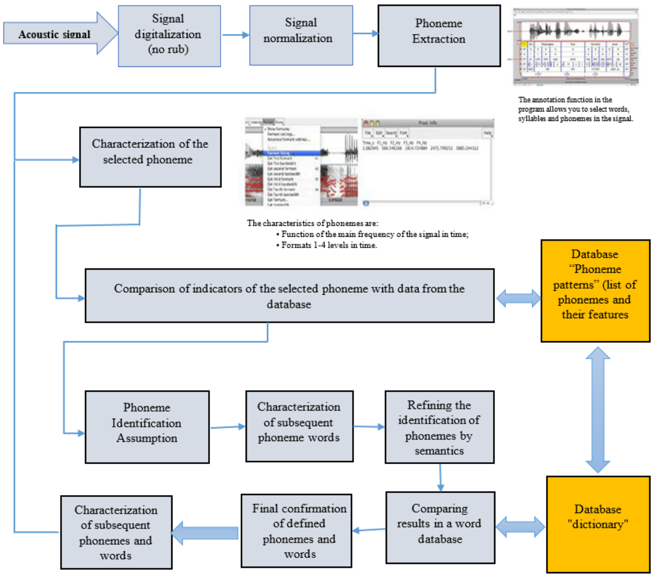

Fig-3. Algorithm for improving the quality of speech signals

In practice, attention is not paid to high frequency ranges in practice, since it is generally accepted that they do not play any significant role in speech recognition and identification. High-frequency ranges are also not included in the encoding of digital speech signals. In order to ensure the compactness of the speech audio file, they themselves will usually find themselves outboard with digital information.

The software package for testing the proposed algorithms was developed in the C ++ programming language. In total, 4 different phrases were taken from 70 people, of which 30 were male voice signals. The speech signal was recorded in the form of a 16-24-bit audio file with a frequency of 16-44 kHz in stereo mode. The time of pronounced sentences varies from 30 to 60 seconds, and the duration of the control signal is 20-60 seconds. Checking the performance of the algorithm was carried out with a different number of people of different sexes. The following table shows the results.

Number of people.	Recognition percentage on standard range	Recognition percentage on extended range
10-20 years old men:		
10	80	92
21-50 years old men:		
20	89	95
10-20 years old women:		
15	73	87
21-50 years old women:		
25	84	96

IV. CONCLUSION

As a result of the hissing sound study, it was revealed that the frequency structure and combinations of formant changes were fundamentally the same for both male and female speakers.

Based on the foregoing, it can be unequivocally established that adequate recognition and identification of speech requires their mandatory study in an extended frequency range and put into practice.

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

This circumstance underlines the fundamental possibility of using pattern recognition methods for identification and recognition of speech signals that yield more correct results than the probabilistic considerations currently used.

REFERENCES

1. Chuchupal V.Ya., Chichagov A.S., Makovkin K.A. Digital filtering of noisy speech signals: Messages on computer software. M.: Computer Center of the Russian Academy of Sciences, 1998.52 p.
2. Scalart P., Filho J.V. Speech enhancement based on a priori signal to noise estimation // IEEE international conference on acoustics, speech and signal processing (ICASSP-96). 1996. P. 629-632.
3. Ephraim Y., Malah D. Speech enhancement using a minimum mean square error short time spectral amplitude estimator // IEEE transactions acoustics, speech and signal processing. 1984. P.1109-1121.
4. Ephraim Y., Malah D. Speech enhancement using a minimum mean square error log spectral amplitude estimator // IEEE transactions acoustics, speech and signal processing. 1985. P. 443-445.Ahmad X. M. Introduction to the digital processing of speech signals: textbook. allowance / X. M. Akhmad, V.F. Zhirkov; Vladim. state Vladimir: Publishing house Vladim. state Univ., 2008. ¬ 192 p. ¬ISBN 5¬89368¬751¬5.
5. B.Wiedecke, V.Khamrayeva, M. Payazov: Analysis of components of speech audio signals in the extended frequency range.–Tashkent: Vestnik TUIT, № 1, 2010.
6. B.Wiedecke, V.Khamrayeva, M. Payazov: Sizzling signals in the process of identification and recognition of speech. – Tashkent: Vestnik TUIT, № 3, 2010.
7. A.K. Alimurodov, R.R. Churakov «Rewiew and classification methods for processing speech signals in the speech recognation systems» Measurement. Monitoring. Control. Control № 2. 2015
8. A.B. Davidov «Hilbert-Huang transformation» 2010.
9. Nina Berend, Stefan Kleiner, Ralf Knöbl: Sprachliche Variabilität des Deutschen und ihre Erfassung mit Methoden der automatischen Spracherkennung. Institut für Deutsche Sprache Mannheim, 2005.