Software implementation of the information model of the peripheral auditory system for assessing speech intelligibility

S A Rozhkov, M A Savinykh, V A Trushin

Department of Information Security, Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russia

E-mail: rastr89@mail.ru

Abstract. The paper considers the features of constructing an information model of the human peripheral auditory system in the LabVIEW software environment. A brief description of the adopted information model proposed for implementation in the software environment is given. Detailed explanations are described for each element of the model made in the LabVIEW environment. The main functionality of the user part of the program is presented. The work is, in fact, staged in nature and offers a sequence of further necessary studies of the developed model.

Keywords - speech intelligibility, peripheral auditory system, information model, LabVIEW environment, transversal filter, amplitude compressor, information security.

1. Introduction

Currently, various methods are used to assess speech intelligibility in various applied tasks (construction acoustics, medicine, speech communications, psychoacoustics, information protection, etc.), which can be divided into subjective and objective. The first ones are based on articulation tests involving teams of speakers and auditors; the second ones are based on measuring the parameters of speech signals and noise, followed by calculating intelligibility based on accepted models of the relationship of these parameters with speech intelligibility [1]. In Russia, the formant method of N. B. Pokrovsky is mainly used [2]. However, this method was developed more than seventy years ago and, of course, does not reflect modern achievements in the field of psychoacoustics and physiology.

In this work, the information and measurement model of the human peripheral auditory system, proposed in [3], is taken as the basis, as the most complete, according to the authors, taking into account the peculiarities of the system functioning.

The purpose of the work is to check the correctness of the proposed model and to optimize it based on computer modeling in the LabVIEW environment.

2. Description of the information model of the human peripheral auditory system.

The functional diagram of the model is shown in Figure I [3] and includes the following elements:
Figure 1. Information model of the peripheral auditory system for assessing intelligibility

\( y(t) = k_{1r} \times x(t); z(t) = k_{2r} \times F[y(t)]; F \) – compression function

1. The primary converter is a microphone. It converts the input acoustic signal into an electrical one. It is important to emphasize that the microphone should be measuring with an intrinsic noise level of about 20 dBA.

2. An amplifier with an uneven frequency response (a rise of 12-15 dB in the region of 3-4 kHz). It takes into account the differences in ear sensitivity at different frequencies, namely, increased sensitivity in the 3-4 kHz region, thereby reflecting the work of the outer ear in converting the input acoustic signal [4].

3. Amplitude compressor (cubic root or quasi-logarithm). It reflects the nonlinear dependence of the movement of the eardrum of the middle ear on the amplitude of sound vibrations [4][5].

4. A transversal filter with bands corresponding to the critical hearing bands shown in Table 1. The most suitable model for dividing the signal into critical bands is a transversal FIR filter with branches on all bands. A set of filters allows to divide the signal by frequency bands for their independent analysis, similarly to different sections of the basilar membrane of the human inner ear [6].

5. Integrators that determine the average signal energy in each critical band.

Table 1. Critical hearing bands (according to Zwicker)

| Band Range | Frequency Range |
|------------|-----------------|
| 1. 1-100 Hz | 1. 630-770 Hz   |
| 2. 100-200 Hz | 2. 770-920 Hz   |
| 3. 200-300 Hz | 3. 920-1080 Hz  |
| 4. 300-400 Hz | 4. 1080-1270 Hz |
| 5. 400-510 Hz | 5. 1270-1480 Hz |
| 6. 510-630 Hz | 6. 1480-1720 Hz |
3. Software implementation of the model.

3.1. Software implementation of the human auditory system model.
The considered model of the human peripheral system is implemented in the LabVIEW software environment. The analyzed signal can be read from a .wav file, or generated by the program itself. The program outputs the values of the levels in each frequency band, as well as graphs of the spectrum of the processed signal after each element of the converter path.

3.1.1. Primary converter
In this implementation primary converter is reduced to reading an audio file.

3.1.2. Amplifier with uneven frequency response
In the first implementation of the model, the shape of the gain region is assumed to be symmetrical (Figure 2). This block of the information model does not assume signal attenuation. Therefore, standard FIR and IIR filters are not suitable for solving this problem, since they inevitably have a delay band with almost complete attenuation of the input signal.

![Figure 2. Frequency response of the amplifier](image)

To synthesize the required amplifier, the method of direct and inverse Fourier transforms is applied: a Fourier transform is performed for the input signal, then the resulting spectrum is multiplied by the gain function of the required shape, after which the inverse Fourier transform is performed for the resulting spectrum. As a result, a signal with an amplification in a given frequency band is obtained.

To implement this approach, we used the basic Fourier transform functions available in LabVIEW. To obtain the shaping signal shown in Figure 2, the Gaussian function was used, since various distributions in natural systems most often have exactly this form. The basic expression for the Gaussian function has the form:

\[ g(x) = ae^{-\frac{(x-b)^2}{2c^2}} \]  

(1)

The function itself has the form of a "bell", which has a value close to zero at the edges. Parameter \( a \) determines the height of the "bell", \( c \) is its width, and \( b \) is the offset along the x-axis.

For the solved problem of obtaining the required frequency response form, it is necessary to add 1 to all values of the Gaussian function to exclude signal attenuation in the entire frequency range. Therefore, for the required amplitude gain, we have:

\[ a = (\text{amplitude} - 1) \]  

(2)

The offset of the function along the abscissa axis should be determined in frequency units and numerically depend on the frequency resolution \( df \) (which, in turn, is determined by the sampling frequency \( df = Fs/N \), where \( N \) is the number of signal points). Therefore, for the offset of the "bell" by the freq shift frequency, we have:
\[ b = df \cdot \text{freq shift}. \]  \hspace{1cm} (3)

The width of the "bell" can be set by the dimensionless coefficient width coef. At the same time, in order for the width not to depend on the number of points \( N \), it is necessary to include this number \( N \) in the coefficient. Therefore

\[ \frac{1}{(2c^2)} = \frac{\text{width coef}}{(N/2)^2}. \] \hspace{1cm} (4)

The final formula for the frequency gain function in a given range has the form:

\[ g(x) = (\text{amplitude} - 1) \cdot e^{-\frac{(x-\text{df} \cdot \text{freq shift})^2 \cdot \text{width coef}}{(N/2)^2}} + 1 \] \hspace{1cm} (5)

As a result of the operation of this module, a set of \( N \) points is obtained that correspond in values to the form of the frequency response of the amplifier shown in Figure 2, up to the parameters. The implementation of this module is shown in Figure 3. Next, the resulting function is applied to the processed signal.

**Figure 3.** Implementation of the Gaussian function in LabVIEW to generate a gain function with the required frequency response

### 3.1.3. Amplitude compressor

In essence, the amplitude compressor, simulating the transformation of an acoustic signal in the middle ear, should work according to a power dependence, because the application of the logarithm function to the input signal is impossible for the following reasons:

- The logarithm is not defined for negative values, while the acoustic signal is bipolar.
- The logarithm is not defined for the zero value.

Thus, it is possible to use the cubic root function, which does not have these disadvantages and is close to the logarithm function in form for argument values greater than 1 (with an appropriate scaling factor). Or the logarithm of the input signal shifted by 1 module with the assignment of the sign of the original signal to it (quasi-logarithm).

In the framework of this study, the second approach is implemented (Figure 4), that is, the quasi-logarithm, which is determined by the formula (6).

\[ y(x) = \lg(|x| + 1) \cdot \text{sign}(x) \] \hspace{1cm} (6)

**Figure 4.** Implementation of a quasi-logarithmic function in LabVIEW
3.1.4. Bandpass filtering

The peculiarity of filtering vibrations in the inner ear is that the part of the basilar membrane responsible for a certain critical band is excited by mechanical vibrations of the corresponding frequency range. At the same time, such a mechanical filter does not have a clear boundary between the transmission area and the delay area. Therefore, the frequencies of neighboring critical bands are partially extinguished. The basilar membrane is located in the cochlea so that along the path of the acoustic wave from the middle ear, first follow the areas that perceive (and extinguish) high-frequency components, and then – low-frequency ones. That is, the critical bands are arranged in the order from the 24th to the 1st. In this regard, the model implements 24 bandpass filters, each of which receives an input signal (parallel filtering), as well as before each bandpass filter (except for the first one corresponding to the 24th critical band) there is a notch filter that filters out the components of the previous band, as well as partially of the current one (sequential filtering, the implementation is shown in Figure 5). You can quickly switch between these two filtering options while the program is running. This will be necessary later, to analyze the impact of different versions of this model block.

In both cases, band-pass FIR filters of the 199th order are used with cut-off frequencies corresponding to the boundary frequencies of the critical bands. In the case of sequential filtering, similar notch filters are also used – of the same type, order and with the same boundary frequencies.

![Figure 5. Implementation of a sequential filter in LabVIEW](image)

3.1.5. Integrators

Integrators allow us to estimate the average signal energy in each of the critical bands, which allows us to proceed to the assessment of formant intelligibility. The assessment is made according to the formula:

\[ I_{cp} = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} z_k^2(t)dt \]  \hspace{1cm} (7)

\( z_k(t) \) is the signal at the output of the k-th filter, the input of which is supplied with the original signal after amplification and compression, and in the case of a sequential filter – after the notch filter of the previous band.

In addition, this module implements the power logarithm for the transition to decibels, according to the formula:

\[ I_{dB} = 10 \log \left( \frac{I_{cp}}{I_0} \right) \]  \hspace{1cm} (8)

\( I_0 \) is the power of the acoustic signal at the threshold of audibility.

Such an integrator, respectively, is placed at the output of each bandpass filter. The implementation is shown in Figure 6.
3.2. The user part of the interface
The program interface allows to quickly changing the overall model by turning on and off various elements, as well as changing the parameters of a particular element. In addition, the interface displays graphs of the input signal shape and the amplitude spectrum before (the initial signal) and after filtering. Figures 7-8 show the appearance of the user interface.

Figure 6. Implementation of the integrator in LabVIEW

Figure 7. Program Settings panel
4. Practical part

By the implementation of this information model in the LabVIEW software environment, it is possible to conduct a fairly extensive comparative analysis by including the above-mentioned model blocks in various configurations and evaluating the results obtained in each of them.

First of all, the results of such an analysis will give an understanding of which modules in the model make a significant contribution to the final value of speech intelligibility. In particular, the division into 24 frequency bands may be unnecessary, since the number of formants is small at the boundaries of the frequency range under study. Also, for an amplifier stage with an uneven frequency response, changing the parameters of the frequency response form itself may not make significant differences in the final result. All other blocks are also subject to verification on the same principle.

The next step is to compare the new model with the existing one by analyzing the same source files using two versions of the program: in the new model mode and in the existing one. After analyzing the results of this comparison, it will be possible to say how appropriate it is to use a new model instead of the current one, namely, whether the results of calculating speech intelligibility will change.

5. Conclusion

The implementation of the functions of the external, middle and inner ear in the software environment will allow the simplest and reliable interpretation of the model of the human auditory system, since it allows changing the composition and parameters of individual elements.

Taking into account the features of the LabVIEW software environment, the future prospects of the program may consist in creating a software and hardware complex based on it, providing real-time speech intelligibility calculation.

6. References

[1] Aldoshina I A, Pritts R 2006 *Muzykal'naya akustika* (SPb.: Kompozitor) 760 p
[2] Horev A A, Carev N V 2017 Method and algorithm for the formation of speech-like noise *Vestnik VGU – VSU Bulletin, ser. Systems analysis and information technology* vol 1 pp 57–67 (In Russ.)
[3] Trushin V A 2019 The analysis of the formant method of speech intelligibility estimation as a method of performing indirect measurements *Science bulletin of the Novosibirsk State Technical University* no. 4 (77) pp 135–146 DOI: 10.17212/1814-1196-2019-4-135-146.

[4] Mareev G O 2012 Modern mathematical models of middle ear (review) *Saratov Journal of Medical Scientific Research* vol 8 № 1 p 96–100

[5] Usanov D A 2012 Lazernye avtodinnye izmereniya parametrov dvizheniya barabannoj pereponki *Rossijskij zhurnal biomekhaniki* no16 №1 pp 8–21

[6] Vargin P S Matematicheskaya model' ulitki vnutrennego uha [Electronic resource]. – URL: http://www.telenavt.ru/downloads/Ulitka%20vnutrennego%20uha%20cheloveka.pdf (date of application: 01.06.2021)