=Speech Melody Meter= (SMM-v2)

SMM-

v2 version from 2022-07-26

USER GUIDE

1. Functionality

The **Speech Melody Meter** (**SMM**) software is designed to evaluate the statistical characteristics of speech melody by constructing and displaying histograms of the frequency distribution of occurrence of discrete values of the pitch frequency (F0) over sufficiently long speech intervals.

SMM implements a method for estimating three important statistical characteristics of the histogram: the Center of «mass» of the distribution – C [Hz], the Band width – B [Hz], and the Asymmetry distribution – A [Hz].

SMM has a wide range of options for choosing linear scales for constructing histograms used in the analysis of speech melody characteristics, namely: the range of analyzed frequencies – F0min and F0max, as well as frequency resolution – dF0. The choice of specific values of these parameters determines the number of histogram components - N.

SMM is implemented as a separate application for Windows (7, 8, 10).

A distinctive feature of the software implementation of SMM is that the discrete values F0 of speech are measured only at intervals of vowels, which makes it possible to identify the most significant parts of the F0 trajectory for perception and at the same time avoid a number of artifacts that occur at intervals of consonants and pauses.

2. Applicationsc

The use of **SMM** can be useful in the following practically important cases:
- In scientific research of the different expressive function of speech melodic: expressing the personal accent, normativity, style and artistry;
- When studying and objectively assessing the degree of emotionality of one's speech or the speech of the interlocutor;
- For individual training in preparing speeches and presentations for self-control and teaching the desired characteristics of speech melody;

- As a means of monitoring speech intonation in the real work conditions of call center operators, radio and television announcers.
- As a means of self-control to eliminate speech defects associated with various physiological disorders and diseases;

3. Getting Started

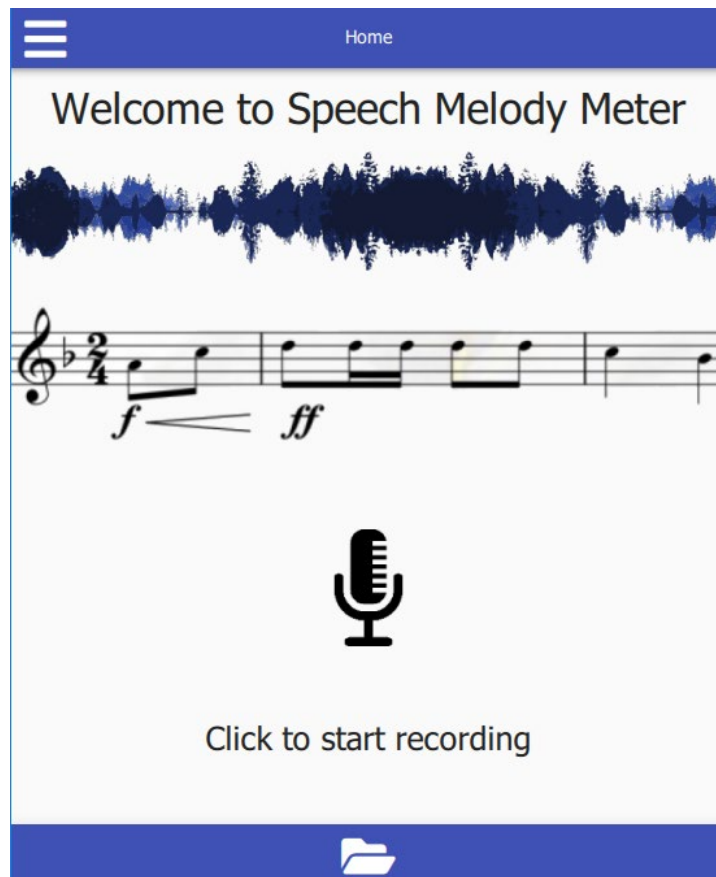The SMM Start Window, which opens after the program is launched, is shown in Figure 1.

Fig. 1 Start Window

Immediately after starting the program, the user is provided with the following options:

1) Recording through a microphone of a sufficiently long segment of the analyzed speech (preferably at least 10 seconds);
2) Calling one of the pre-prepared test audio files stored in the **"data/tests"** folder.

The first option is implemented by clicking on the microphone icon:

After the record icon appears:

speech recording starts. You must click this icon again to end the recording.

The spoken speech signal is recorded in the **"data/records"** folder and prepared for further processing and displaying the results.

The second option is implemented by clicking on the icon:

to call pre-prepared test audio files from the **"data/tests"** folder (see section 6).

4. Graphical display of the results

After the analyzed speech signal is entered from the microphone or from the "data/tests" folder, the Main SMM window is opened (see Fig. 2) and the results of the histogram calculation and 3 of its statistical characteristics are displayed.
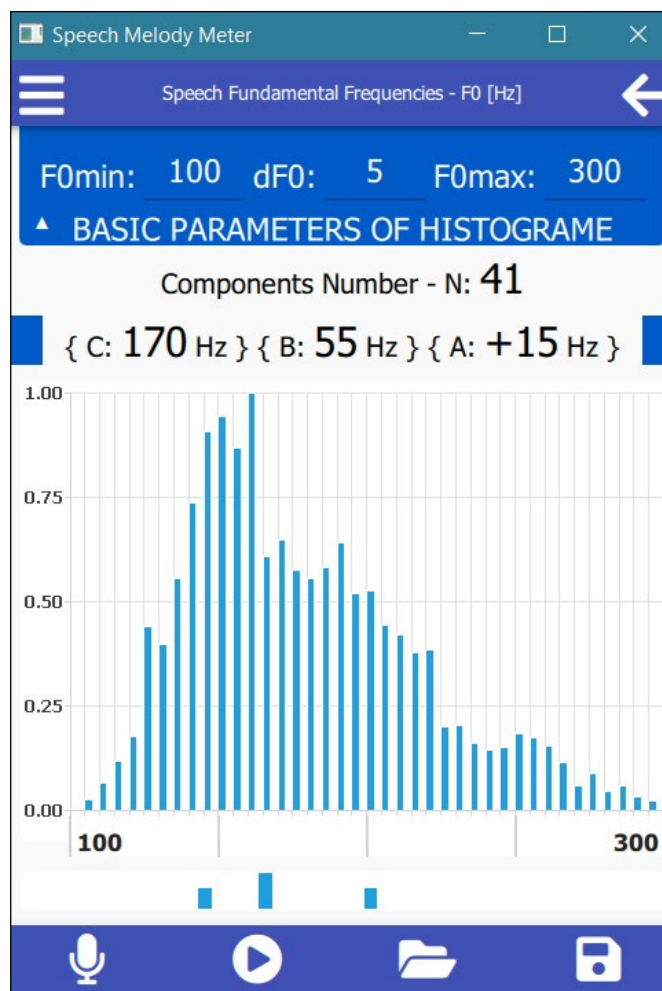
Fig.2 Main Window

The central part of the Main window (Fig. 2) shows the histogram of the analyzed segment of the speech signal. The vertical axis of the histogram shows the normalized values of the frequency of occurrence of discrete values F0. The horizontal axis shows

the F0 scale indicating the chosen by user diapason of pitch frequencies analysis (from 100 Hz to 300 Hz).

Directly below the scale, three vertical marks are shown, indicating the position of the center of the "mass" of the histogram - C (longer one) and two shorter ones, defining the band width - B and asymmetry - A of the distribution.

In the upper half of the main window, the blue rectangle displays the values of the histogram parameters selected by the user: F0min = 100 Hz, dF0 = 5 Hz, F0max = 300 Hz, that necessary for calculating and constructing the histogram of the segment of the speech analyzed. Below are the results of calculating the statistical characteristics of the presented in fig. 2 histograms, namely:
- Number of components - N: 41;
- Center of "mass" - C: 170 Hz;
- Effective bandwidth - B: 55 Hz;
- Distribution asymmetry - A: +15 Hz.

Note that, as a result of the calculation, the asymmetry can take a positive value (+) A, if the histogram is dominated by high-frequency components, and a negative value – (-) A, if it is dominated by low-frequency one.

An important feature of the program is that the User, in accordance with the task he is solving, is given the opportunity to quickly change the corresponding numerical values F0min, F0max, dF0. These parameters determine the frequency range and frequency resolution of the analysis, as well as the number of histogram components - N.

After changing the values of one or more initial parameters the icon appears:

By activating this icon, when analyzing the same fragment of the speech signal, the type of the histogram, as well as its statistical characteristics, is recalculated in accordance with the new values of the parameters.

An important feature of the SMM is the possibility for the user to visually compare of two histograms received from different sources of speech. At the same time, it is provided the calculation and displaying of the statistical characteristics of both signals: (C1 - C2), (B1 - B2), (A1 - A2).

The possibility of obtaining two histograms is provided as follows:
- at the first signal input, a blue histogram is displayed, its statistical characteristics are calculated and displayed.
- when another signal is input, new blue histogram is displayed and its statistical characteristics are displayed. At the same time, the old histogram and its characteristics are preserved, but it changes color from blue to green.

The usefulness of visual comparison of histograms and their statistical characteristics one can see on the SMM Main Window (see Fig. 3 and Fig. 4).

Fig. 3 shows the histograms of reading the same story by two professional male speakers, whose are distinguished by a peculiar intonation expressiveness.
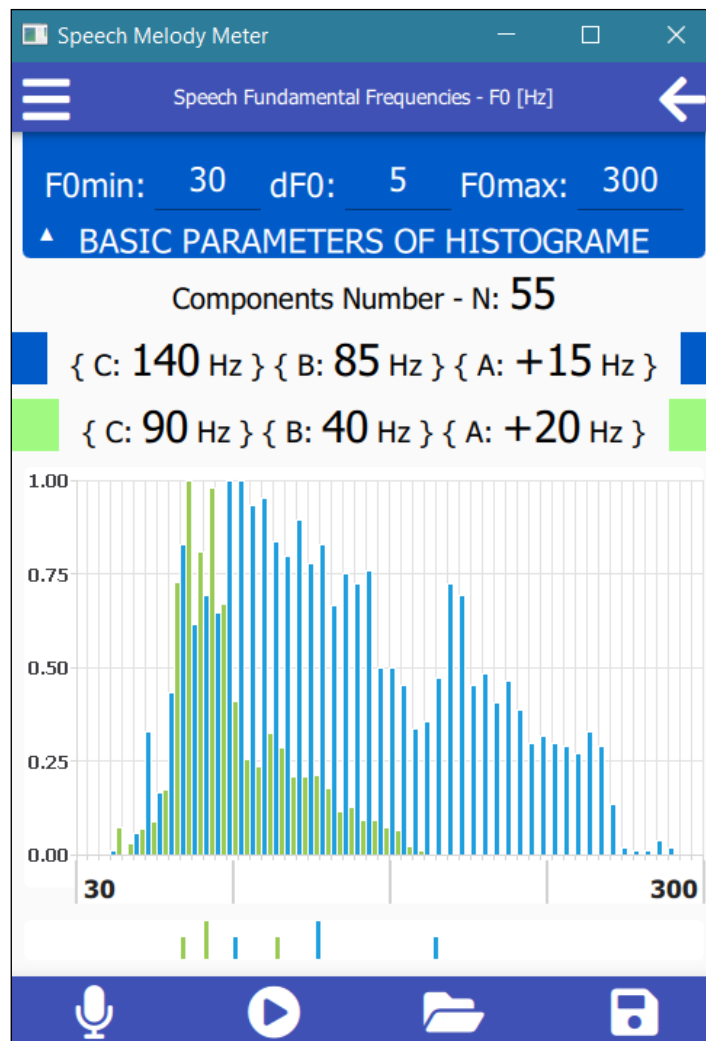
Fig. 3. The view of SMM Main Window for the performances of two different professional male speakers (green and blue)

On fig. 4 histograms of two speeches by the same TV male presenter with neutral (blue) and angry (green) emotion are shown.
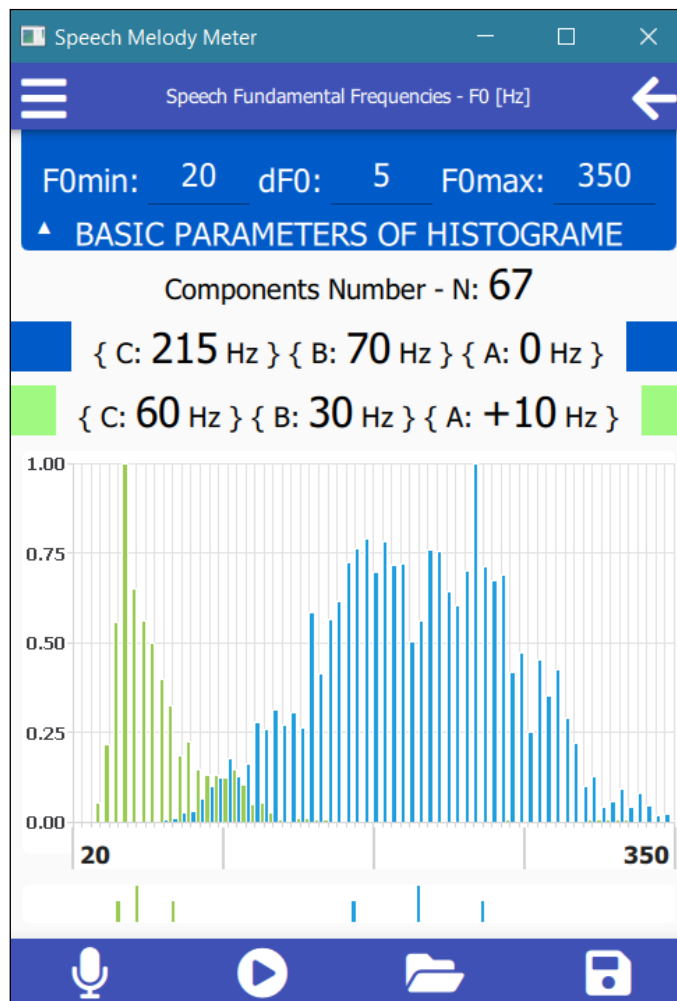
Fig. 4. The view of SMM Main Window in the TV presenter's speech with neutral (green) and angry (blue) emotion

5. Auxiliary icons for displaying results

In the lower blue bar of the main window, in addition to the icons for inputting a signal from a microphone and calling test audio files from the "data/tests" folder there are also two other icons.

Using icon

,

ou can listen to the speech recorded through the microphone or called from the "data/tests" file.
Using the second icon,

the results of measuring the characteristics of the histogram of the analyzed speech are recorded in the "results" file (see Fig. 6).

In the upper blue bar of the main window on the left is an icon, when clicked, an additional information window opens (see Fig. 5).,
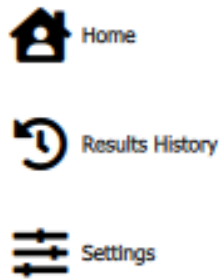
Fig. 5. Information window

By choosing an icon: the contents of the "results" file are called (see Figure 6).

Fig. 6. The contents of the "results" file

By choosing an icon and its activation window is called.

,

It designed to set the threshold HTh of the minimum value of the histogram components by the user.

On fig. 7 is an illustration of the appearance of the histogram when different threshold values are set.
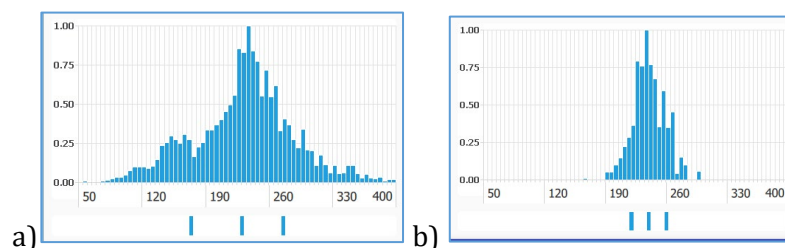
Fig. 7. Illustration of the type of histogram for different threshold values:
a) HTh = 0.0; b) HTh = 0.3

ATTENTION! The internal information in Settings called up by ticking the top right box "Show Advanced" is for the developer only. Changing its contents by the user is highly undesirable.

6. Test Audio Files

As examples of the functioning of SMM, the "tests" folder contains samples of natural speech from various sources (see Fig. 8).
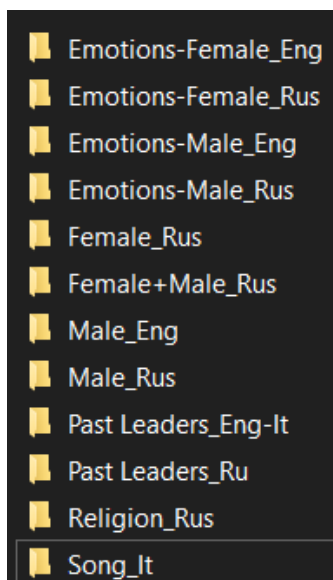

Fig. 8. The "tests" folder content

The "tests" folder contains examples:
- speech in English and Russian,
- calm and emotional speech,
- high and low female voices,
- high and low male voices,
- voices of famous leaders of the past,
- religious prayer and preaching,
- opera aria for tenor and soprano.

The user, based on his own scientific or practical interests, is given the opportunity to record an unlimited number of speech audio files from various sources. To do this, he needs to open the "**SpeechMelodyMeterBin**" folder downloaded from the site, select the "**data**" folder, and then the "**tests**" folder. The contents of the "**tests**" folder can be updated and voiced using the available tools (eg Sound Forge).

Keep in mind that this version of SMM provides the ability to analyze only wav, mono, 8 kHz signals. Any other formats should be adapted to this format using suitable software tools.

CONCLUDING REMARKS

We encourage users to check our site from time to time for useful SMM updates.
To learn more about the theoretical foundations and code of SMM development see additional information posted on this site.