

=Speech Melody Meter= (SMM-v2)

Версия SMM-v2 от 2022-07-26

РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ

1. Функциональное назначение

Программа **Speech Melody Meter (SMM)** предназначена для оценки статистических характеристик мелодики речи путём построения и отображения гистограмм распределения частот встречаемости дискретных значений частоты основного тона (F_0) на достаточно длительных интервалах речи.

В SMM реализован метод оценки трех интегральных характеристик гистограммы: центра «масс» распределения – C [Hz], эффективной полосы – B [Hz], асимметрии распределения – A [Hz].

В SMM предусмотрена возможность широкого выбора параметров линейных шкал гистограммы, а именно: диапазона анализируемых частот – F_{0min} и F_{0max} , частотного разрешения – dF_0 , числа компонент гистограммы – N_c .

SMM реализован как программное приложение для Windows (7, 8, 10).

Отличительной особенностью программной реализации SMM является то, что дискретные значения F_0 речи измеряются только на интервалах гласных звуков, что позволяет выделить наиболее значимые для восприятия участки траектории F_0 и в то же время избежать ряда артефактов, возникающих на интервалах согласных и пауз.

2. Области применения

Использование **SMM** может быть полезным в следующих практически важных случаях:

- ✓ В научных исследованиях различных выразительных функций мелодики: персонального акцента, нормативности, стиля и артистичности речи;
- ✓ При изучении и объективной оценке степени эмоциональности своей речи или речи собеседника;
- ✓ В качестве тренажера нормативных характеристик мелодики речи при подготовке выступлений и презентаций.

- ✓ Как средство мониторинга характеристик мелодики речи в реальных условиях работы операторов колл-центра, дикторов радио и телевидения.
- ✓ В качестве средства самоконтроля для устранения дефектов речи, связанных с различными физиологическими нарушениями и болезнями;

3. Начало работы

Стартовое окно **SMM**, открывающееся после запуска программы, показано на Рис. 1.

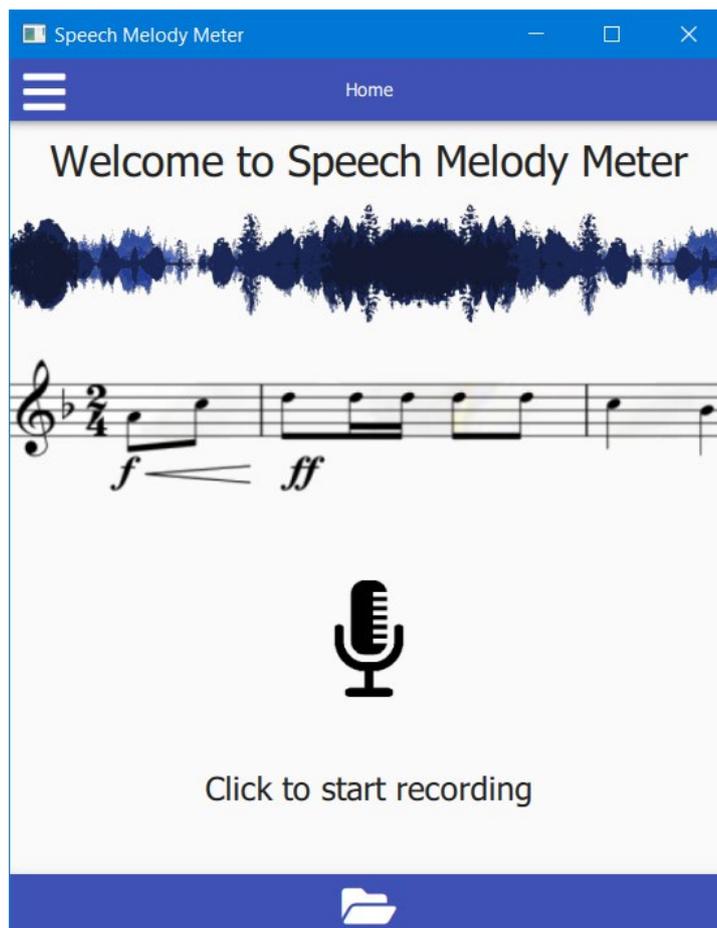


Рис. 1. Стартовое окно SMM

Сразу после запуска программы пользователю предоставляются следующие возможности:

- 1) Запись через микрофон достаточно длительного отрезка анализируемой речи (желательно не менее 10 сек);
- 2) Вызов одного из заранее подготовленных тестовых аудио файлов, хранящихся в папке “**data-tests**”.

Первый вариант реализуется путём нажатия на значок микрофона



и, после появления значка записи:



начинается запись произносимой речи. Для завершения записи необходимо повторное нажатие этого значка.

Произнесённый речевой сигнал записывается в папку “**data/records**” и подготавливается для дальнейшей обработки и отображение результатов.

Второй вариант реализуется путём нажатия на значок,



предназначенный для вызова из папки “**data/tests**” заранее подготовленных тестовых аудио файлов (см. раздел 6).

4. Графическое отображение результатов

После того, как анализируемый речевой сигнал введен с микрофона или из папки “**data/tests**”, осуществляется открытие **Главного окна SMM** (см. рис.2) с отображением результатов расчета гистограммы и 3-х её статистических характеристик.

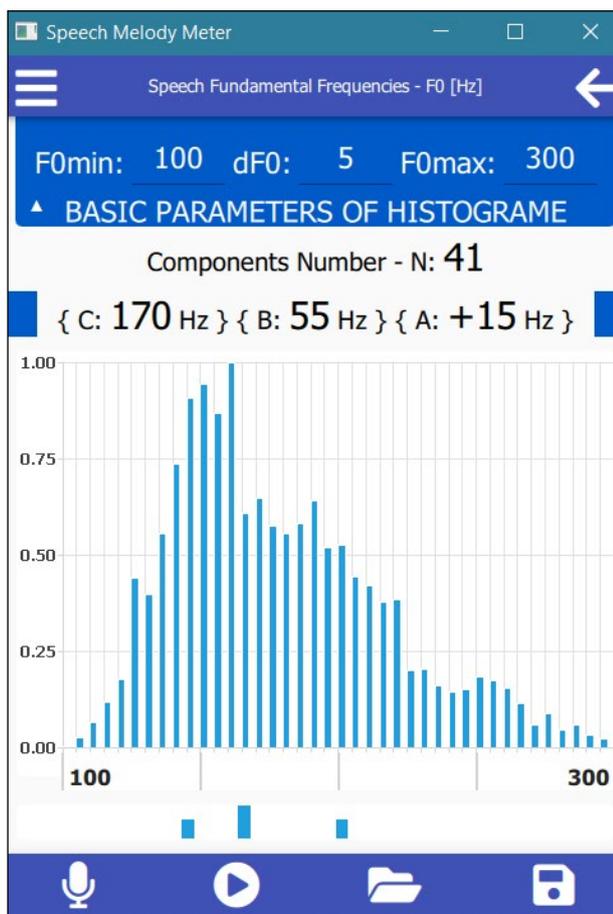


Рис. 2. Главное окно SMM

В центральной части Главного окна (рис. 2) представлена гистограмма анализируемого сегмента речевого сигнала. По вертикальной оси гистограммы отложены нормированные значения частоты появления дискретных значений F_0 . По горизонтальной оси отложена шкала F_0 , указывающая выбранный пользователем диапазон анализа частот основного тона (от 100 Гц до 300 Гц).

Непосредственно под шкалой нанесены три вертикальные метки, обозначающие положение центра «массы» гистограммы – С (более длинная) и две более коротких, определяющих ширину полосы – В и асимметрию – А распределения.

В верхней половине основного окна в синем прямоугольнике отображаются выбранные пользователем значения параметров гистограммы: $F_{0min} = 100$ Гц, $dF_0 = 5$ Гц, $F_{0max} = 300$ Гц, необходимые для расчета и построения гистограммы отрезка анализируемой речи. Ниже приведены результаты расчета статистических характеристик представленной на рис. 2 гистограммы, а именно:

- Количество компонентов - N: 41
- Центр «массы» - С: 170 Гц
- Эффективная ширина В: 55 Гц
- Асимметрия - А: +15 Гц

Отметим, что в результате расчета асимметрия может принимать положительное значение - (+) А, если на гистограмме преобладают высокочастотные составляющие, и отрицательное значение - (-) А, если на ней преобладают низкочастотные компоненты.

Важной особенностью программы является то, что Пользователю, в соответствии с решаемой им задачей, предоставлена возможность оперативного изменения соответствующих численных значений F_{0min} , F_{0max} , dF_0 . Эти параметры определяют частотный диапазон и частотное разрешение анализа, а также количество компонент гистограммы – N.

При изменении значений одного или нескольких исходных параметров появляется значок:



После активации этого значка происходит перерасчёт вида гистограммы, а также её интегральных характеристик, в соответствии с новыми значениями параметров.

Важной особенностью SMM является возможность визуального сравнения пользователем двух гистограмм, полученных от разных источников речи. При этом обеспечивается расчет и отображение статистических характеристик обоих сигналов: (С1 - С2), (В1 - В2), (А1 - А2).

Возможность получения двух гистограмм обеспечивается следующим образом:

- при первом входе сигнала высвечивается синяя гистограмма, вычисляются и выводятся ее статистические характеристики.

- при поступлении другого сигнала отображается обновленная синяя гистограмма и ее статистические характеристики. При этом старая гистограмма и ее характеристики сохраняются, но она меняет цвет с синего на зеленый.

Полезность визуального сравнения гистограмм и их статистических характеристик можно увидеть в Главном окне SMM (см. рис. 3 и рис. 4).

На рис. 3 представлены гистограммы чтения одного и того же рассказа двумя профессиональными дикторами-мужчинами, отличающимися своеобразной интонационной выразительностью.

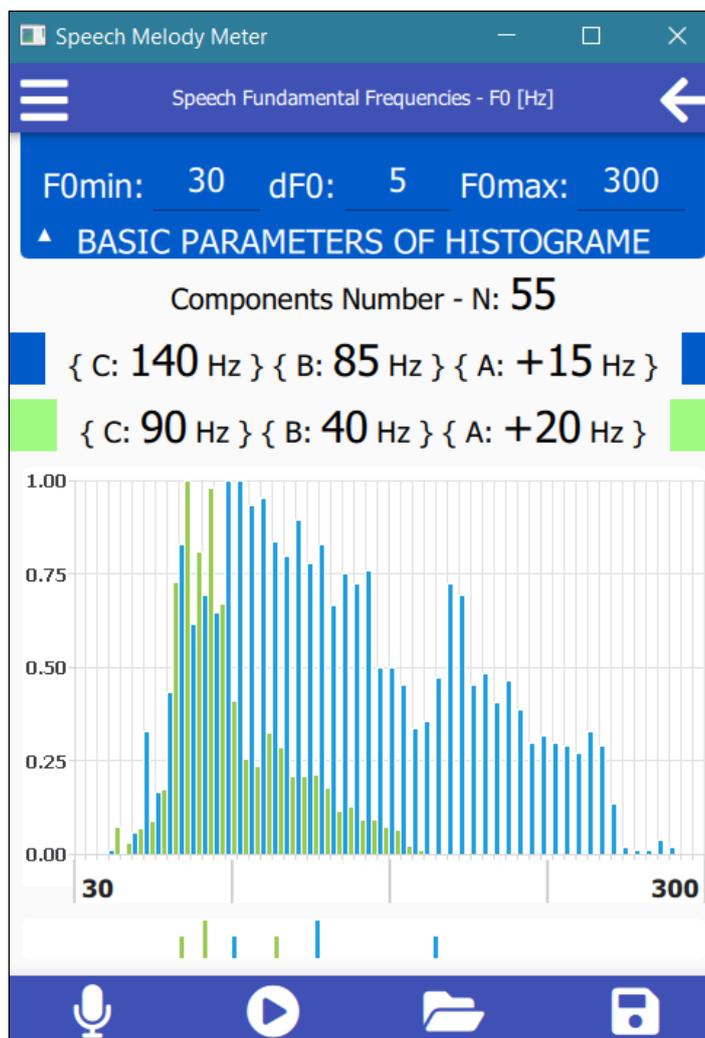


Рис. 3 Вид главного окна SMM при чтении одного и того же рассказа двумя профессиональными дикторами-мужчинами (зеленые и синие)

На рис. показаны 4 гистограммы двух выступлений одного и того же телеведущего с нейтральной (синий) и гневной (зеленый) эмоциями.

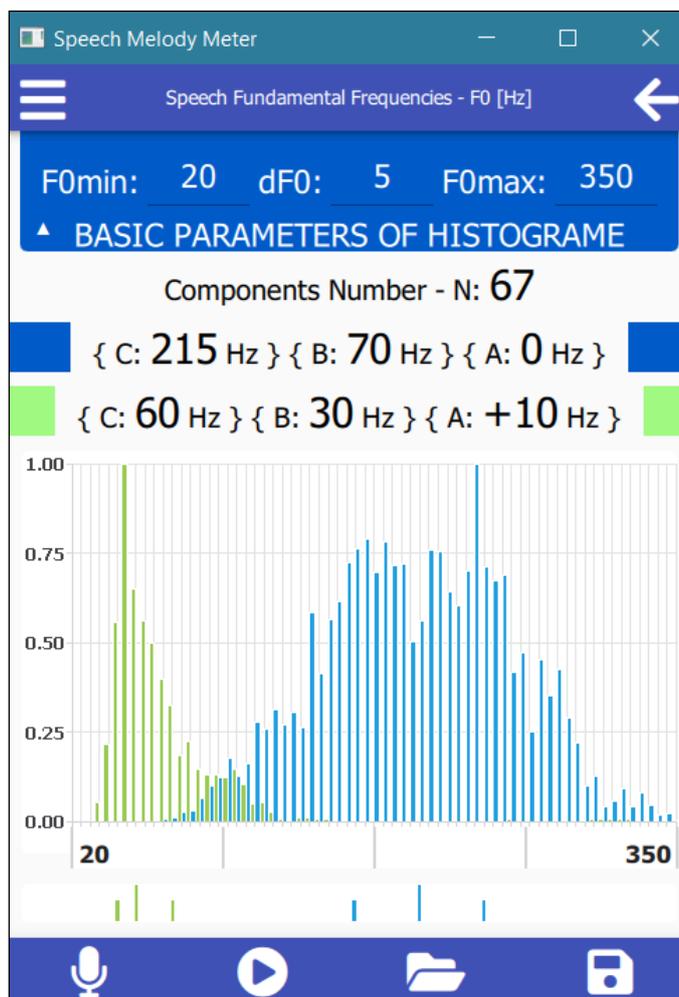


Рис. 4 Вид главного окна СММ в выступлении телеведущего с нейтральной (зеленая) и гневной (синяя) эмоциями

5. Вспомогательные значки для отображения результатов

В нижней синей полоске главного окна, кроме значка вызова тестовых аудио файлов из папки “**data/tests**”, находится также два других значка.

Используя значок , можно прослушать речь, записанную через микрофон или вызванную из файла “**data/tests**”.

С помощью второго значка  осуществляется запись результатов измерения характеристик гистограммы анализируемой речи в файл “**results**” (см. рис. 6).

В верхней синей полоске главного окна слева находится значок,



при нажатии на который открывается дополнительное информационное окно (см. рис. 5),

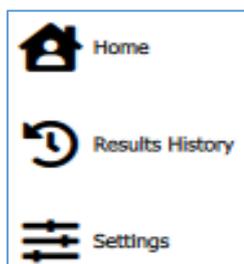
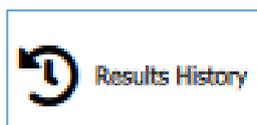


Рис. 5. Информационное окно



Путём выбора значка: вызывается содержимое файла “results” (см. рис. 6).

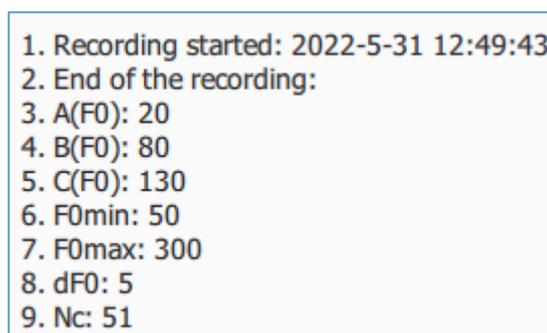
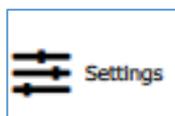
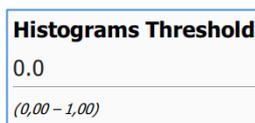


Рис. 6. Содержимое файла “results”



При выборе значка и его активации вызывается окно



предназначенное для установки пользователем порога минимального значения компонент гистограммы HTh.

На рис. 7 дана иллюстрация вида гистограммы при установке различных значения порога.

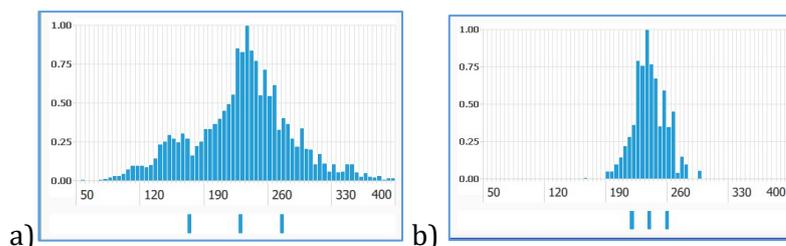


Рис. 7. Иллюстрация вида гистограммы при различных значения порога:
a) HTh = 0.0; b) HTh = 0.3

ВНИМАНИЕ! Внутренняя информация, вызываемая пометкой в правом верхнем квадрате “*Show Advanced*”, предназначена только для разработчика. Изменения её содержимого пользователем крайне нежелательно.

6. Тестовые аудио файлы

В качестве примеров функционирования SMM в папке “**tests**” размещены образцы естественной речи из различных источников (см. рис. 8).

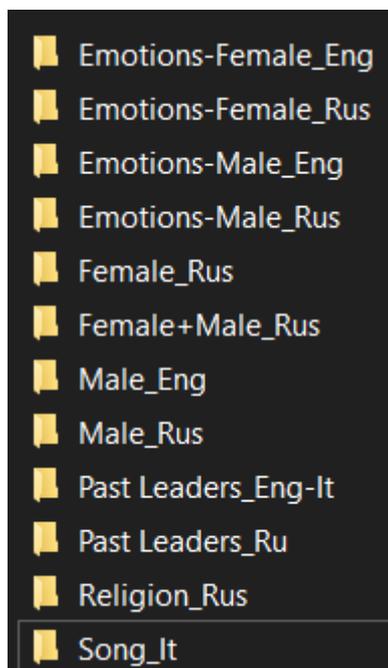


Рис. 8. Содержимое папки “**tests**”

Папка “**tests**” содержит примеры:

- на английском и русском языках,
- спокойной и эмоциональной речи,
- высоких и низких женских голосов,
- высоких и низких мужских голосов,
- голоса известных лидеров прошлого,
- религиозная молитва и проповедь,
- оперная ария в исполнении тенора и сопрано

Пользователю, исходя из его научных или практических интересов, предоставляется возможность записи неограниченного количества собственных речевых аудио файлов. Для этого ему необходимо открыть скачанную с сайта папку “**SpeechMelodyMeterBin**“, выбрать папку “**data**“, а затем папку “**tests**“. Содержимое папки “**tests**” может пополняться и озвучиваться с использованием имеющихся средств (например, **Sound Forge**).

Следует иметь в виду, что в настоящей версии **SMM** предусмотрена возможность анализа только для сигналов в формате: **wav, mono, 8 kHz**. Все возможные другие форматы следует адаптировать под этот формат, используя подходящие программные средства.

ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Мы рекомендуем пользователям время от времени заглядывать на наш сайт, в котором постоянно могут появляться полезные обновления SMM.

Узнать больше о теоретических основах и программном коде разработки SMM: см. дополнительную информацию, размещаемую на этом сайте.