

=VoiceEmoMeter= (VEM)

Версия VEM от 2023-07-18

РУКОВОДСТВО ПОЛЬЗОВАТЕЛЯ

1. Назначение

Программа **VoiceEmoMertter (VEM)** предназначена для измерения **степени эмоциональности** голоса человека на шкале от 0° до 100°:

- Низкая степень от 0° до 30° - «*Вы спокойны, расслаблены и контролируете ситуацию*» (размышления, воспоминания, устное чтение и др.);
- Средняя степень от 30° до 70° - «*Вы активны и уверенно держите себя в руках*» (диалог, выступление, лекция и др.);
- Высокая степень от 70° до 100° - «*Вы возбуждены и не способны контролировать ситуацию*» (злость, истерия, агрессия и др.).

VEM приспособлен для анализа эмоциональности мужских и женских голосов. Измерение степени эмоциональности голоса может быть осуществлено, как для записанного непосредственно с микрофона речевого фрагмента, так и для записанных заранее тестовых аудио файлов из различных источников.

VEM реализован как отдельное приложение для **Windows (7, 8, 10)**.

2. Области применения

VEM может быть востребован для контроля степени эмоциональности голоса в телефонии, радио и ТВ вещании, в обучении, медицине, психиатрии, а также в других общественных сферах жизни.

VEM может быть полезен в повседневной семейной и общественной жизни для мониторинга эмоциональной атмосферы в доме или на работе.

VEM может стать персональным средством для людей различных профессий, стремящихся к повышению EQ своего, так называемого, «эмоционального интеллекта».

3. Метод измерения

Для измерения степени эмоциональности голоса используются определенные параметры гистограммы распределения частоты основного тона (F0) анализируемых звуковых образцов. Измерения проводятся отдельно в диапазонах изменений F0 для мужских и женских голосов.

На рис. 1а представлены две F0-гистограммы образцов мужского голоса в диапазоне 60 – 330 Гц: с низкой степенью эмоциональности (голубая кривая) и с высокой степенью (коричневая), а на 1б – (в диапазоне 130 – 400 Гц) аналогичные гистограммы женского голоса.

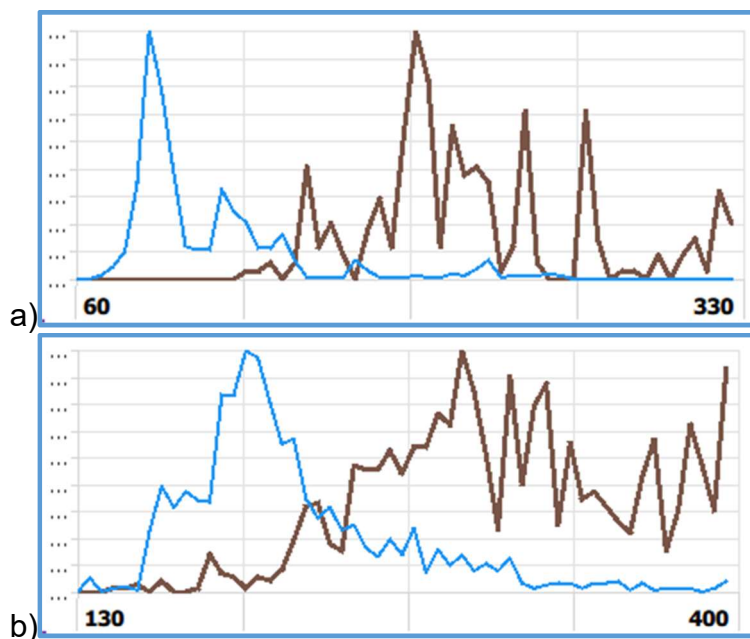


Рис. 1. F0 – гистограммы мужского и женского голосов

Визуальное сравнение F0-гистограмм с низкой и высокой степенью эмоциональности, представленных на рис. 1а и 1б, наглядно показывает их существенное различие, как для мужского, так и для женского голоса.

Для оценки интегральных различий F0-гистограмм с низкой и высокой степенью эмоциональности используются следующие два параметра:

- центр “тяжести” гистограммы – С,
- эффективная “ширина” гистограммы – В.

Численные значения обоих параметров изменяются от самых малых значений – C_{min} , B_{min} - для голоса с низкой степенью эмоциональности, до самых больших значений – C_{max} , B_{max} - для голоса с высокой степенью эмоциональности.

Степень эмоциональности E^o голоса определяются в шкале «градусов» от 0^o до 100^o по формуле:

$$E^o = \{[(C+B) - (C_{min} + B_{min})] / [(C_{max} + B_{max}) - (C_{min} + B_{min})]\} \times 100^o. \quad (1)$$

Здесь:

С, В – значения этих параметров, полученные в процессе измерения фактической эмоциональности голоса конкретной личности в конкретной ситуации.

C_{min} , V_{min} – значения параметров базового эмоционально нейтрального персонального образца голоса (с низкой степенью эмоциональности), полученного на подготовительном этапе (см. п. 4).

C_{max} , V_{max} – ожидаемые значения параметров эмоционально возбужденного персонального образца голоса (с высокой степенью эмоциональности). Расчет численных значений этих параметров осуществляется по формулам:

$$\begin{aligned} C_{max} &= C_{min} * dC, \\ V_{max} &= V_{min} * dV. \end{aligned} \quad (2)$$

Здесь множители dC и dV определяются на основе статистических исследований. Они всегда больше единицы и представлены в разделе настроек отдельно для мужских и женских голосов.

4. Подготовка к измерению E^o

Стартовое окно **VEM**, открывающееся после запуска программы, показано на Рис. 2.

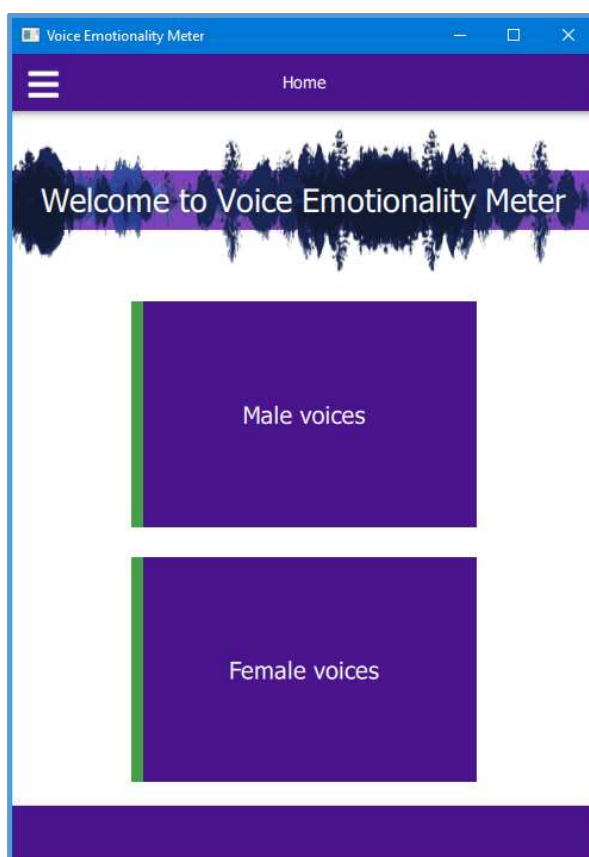


Рис. 2. Стартовое окно VEM

После запуска программы пользователю предлагается выбор гендерной принадлежности тестируемого голоса: мужской или женский.

4.1. После указания гендерной принадлежности, например, **Male voice**, открывается подготовительное окно - Рис. 3а. Аналогичное окно открывается при выборе опции **Female voice** - Рис. 3б.

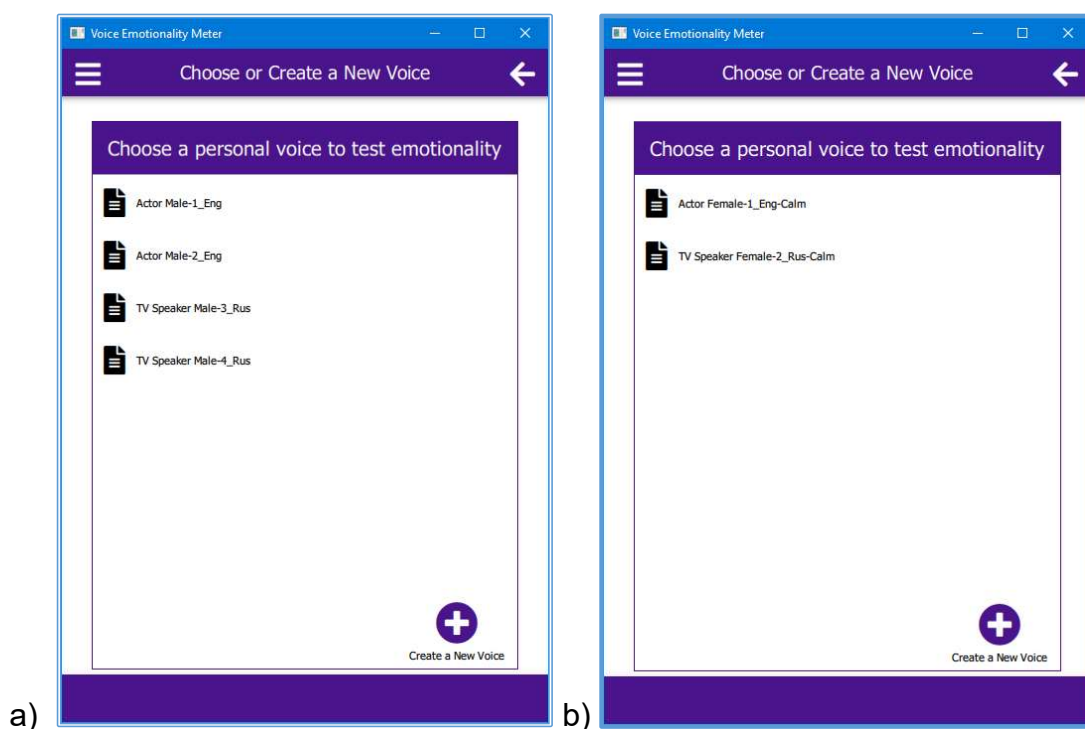


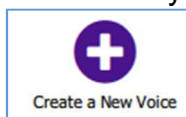
Рис.3. Подготовительное окно (а - мужской, б - женский голос)

Это окно обеспечивает выбор из предварительно созданных базовых образцов голоса, а также дает возможность создания новых персональных базовых образцов. Подготовительные режимы работы для мужского и женского голосов идентичны.

В качестве примеров в папке *data/patterns/male* VEM хранятся образцы мужских голосов: **Actor Male-1**, **Actor Male-2** из “*The Ryerson Audio-Visual Database of Emotional Speech and Song*”, а также образцы голосов известных дикторов русского ТВ – **TV Speaker Male-3**, **TV Speaker Male-4** (см. рис. 3а).

Важно отметить, что для измерения степени эмоциональности голоса конкретной личности пользователю необходимо заранее создать или позаимствовать из других известных источников его **базовый образец голоса с низкой степенью эмоциональности**.

4.2. Для создания нового эмоционально нейтрального персонального голоса, соответствующего низкой степени эмоциональности, используется значок



(см. Рис. 3).

При его активации открывается дополнительное окно (Рис. 4), позволяющее записать звуковой образец голоса личности и его персональное имя, а также сохранить звук в wav-формате.

Рис. 4. Дополнительное окно

4.3. Оперативная запись звукового образца персонального голоса человека осуществляется через микрофон путём нажатия на значок микрофона:



Затем, после появления значка:



начинается запись голосового фрагмента.

Для завершения записи необходимо повторное нажатие этого значка.

Исполняемый голосовой фрагмент записывается в папку **“data/patterns/male(female)”** и отображается на подготовительном окне (см. Рис. 5, **Boris_12-07-23**). В дальнейшем он используется в качестве базового эмоционально нейтрального образца при измерении степени эмоциональности голоса этого человека в различных эмоциональных состояниях.

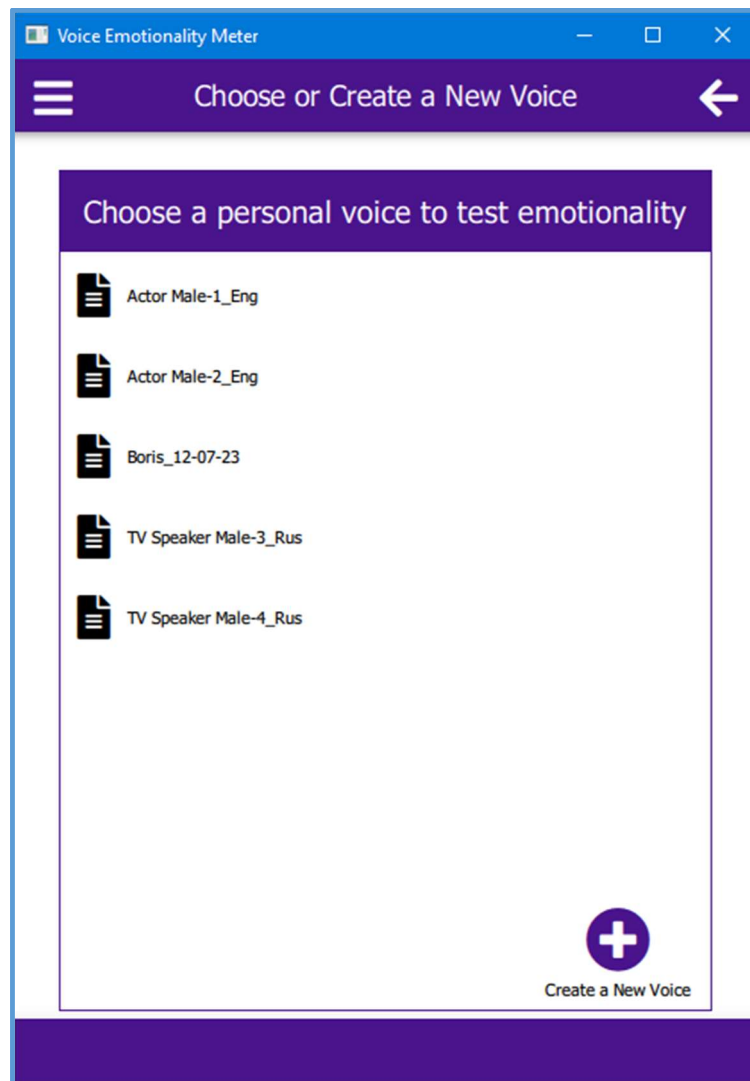


Рис. 5. Отображение результата записи образцов

Описанная процедура может быть продолжена многократно для записи образцов голоса других личностей.

Длительность записи должна быть не менее 10 с.

5. Рабочий режим измерения E^o

После того, как создан набор персональных звуковых образцов (см. п. 4.2, 4.3), можно перейти в рабочий режим измерения E^o .

При выборе одного из образцов (см. Рис.5) (например, базового образца голоса "*Actor Male-2_Eng*") открывается начальное рабочее окно (см. Рис. 6) с изображением гистограммы распределения F0 базового образца.

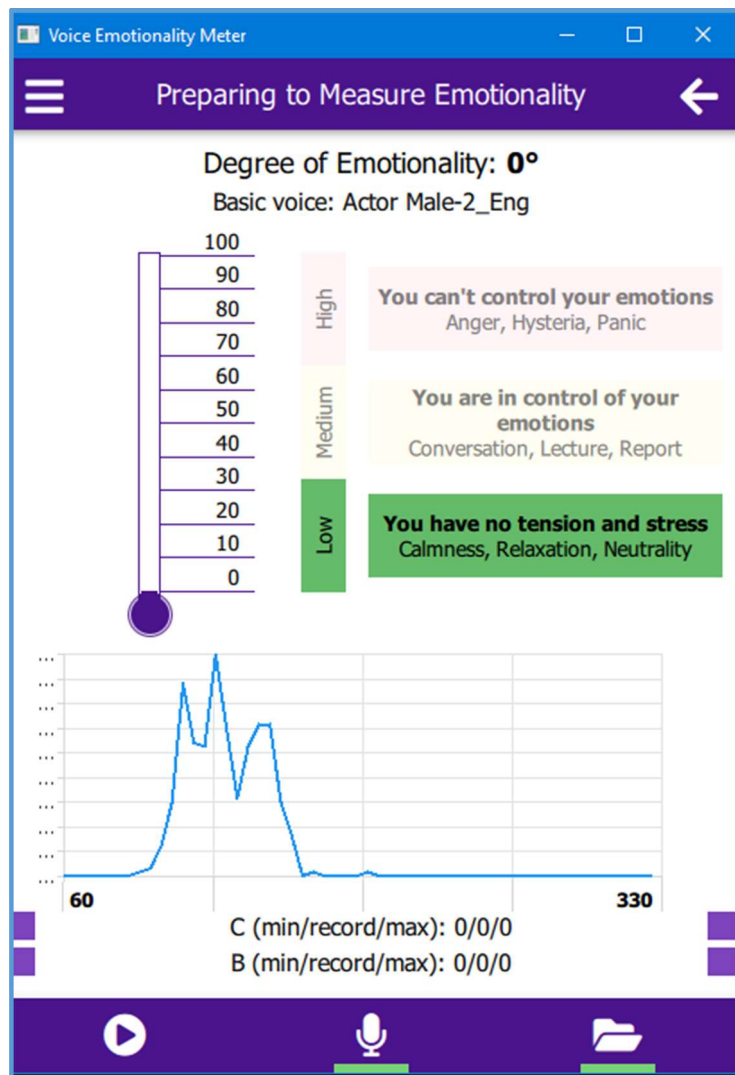



Рис. 6. Начальное рабочее окно VEM

Затем, используя значок , из папки **“data/tests”** вызывается образец этого же голоса – **“Actor Male-2_Eng-Angry”** – с явно выраженной эмоцией гнева. В результате появляется рабочее окно VEM (см. Рис. 7), отображающее результаты измерения степени эмоциональности E^o этого голоса.

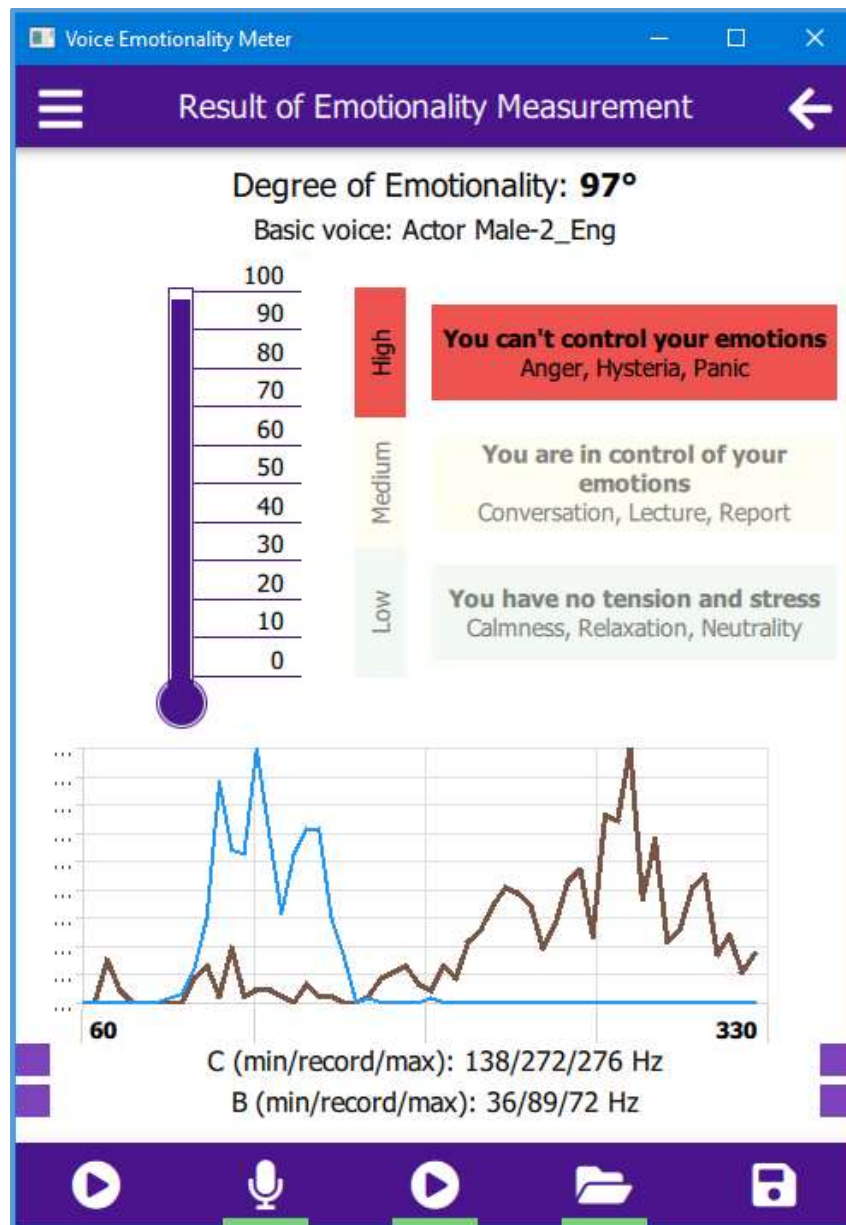


Рис. 7. Конечное рабочее окно (*basic voice Actor Male-2*)

В нижней части рабочего окна голубым цветом изображена гистограмма распределения значений F_0 базового голоса, а коричневым цветом - голоса с эмоцией гнева. Ниже изображения гистограмм приведены соответствующие им значения параметров C и B .

В верхней части рабочего окна показано численное значение измеренной степени эмоциональности гнева – $E^o = 97^\circ$.

В средней части окна наглядно представлена шкала в виде градусника, показывающего измеренную степень эмоциональности.

Справа от шкалы даны надписи, поясняющие ситуативные условия для возникновения каждой из трех обобщенных степеней эмоциональности – низкой, средней и высокой.

Процедура измерения E_o для образцов голоса из “*The Ryerson Audio-Visual Database*”, описанная выше, остается прежней и для случая, когда голосовые образцы конкретной личности записываются через микрофон.

На Рис. 8 представлен результат измерения Ео при выразительном прочтении отрывка стихотворения А.С. Пушкина (базовый голос Борис_12-7-23, нейтральное прочтение того же отрывка).

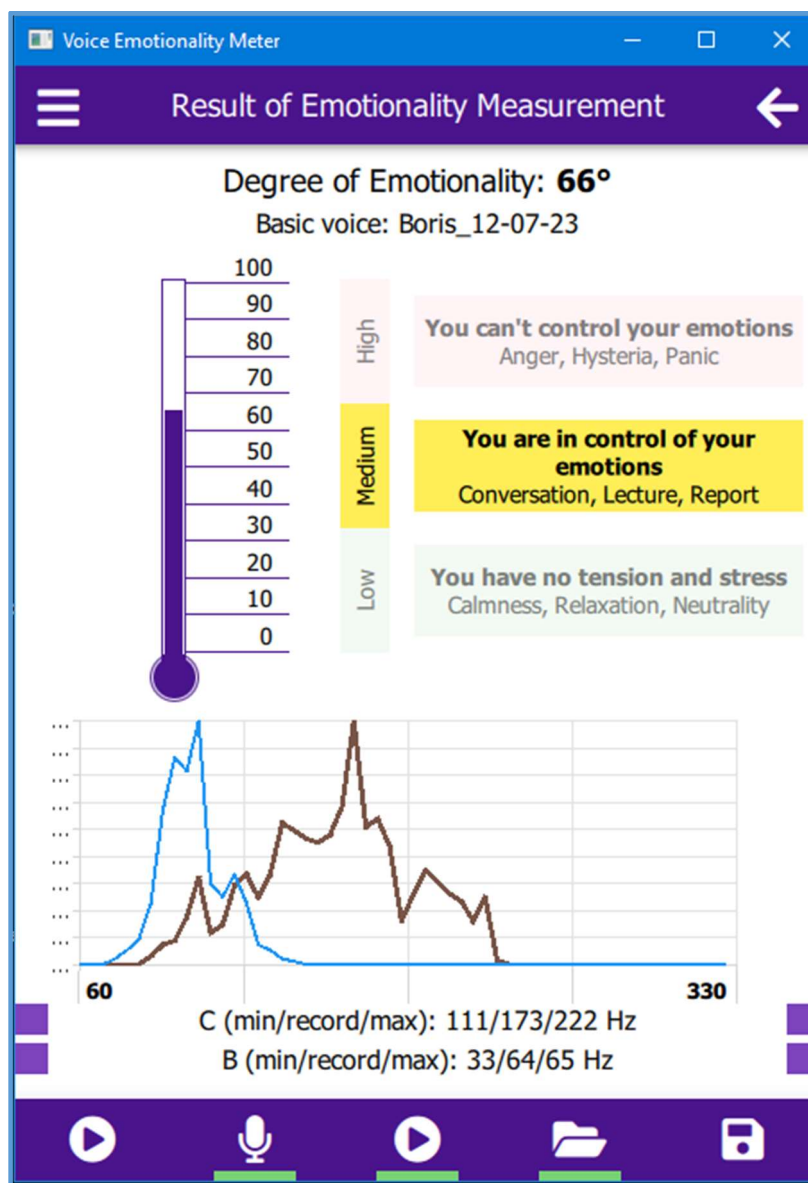






Рис. 8. Конечное рабочее окно (*basic voice Boris_12-7-23*)


6. Вспомогательные значки

- 6.1. В нижней синей полоске на Рис. 6 - 8 значок  используется для прослушивания базового голоса,
- а значок  на Рис. 7,8 - для прослушивания голоса, степень эмоциональности которого измеряется.

6.2. В нижней синей полоске рабочего окна на Рис. 7,8 находится значок микрофона , используемый при записи тестового аудио файла.

6.3. Значок  используется для вызова тестовых файлов из папки “data/tests”.

6.4. С использованием значка  осуществляется сохранение результатов тестирования.

6.5. В верхней синей полоске рабочего окна слева находится значок , при нажатии на который открывается дополнительное информационное окно (см. Рис. 9),

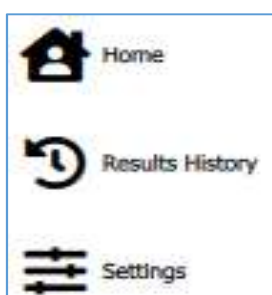
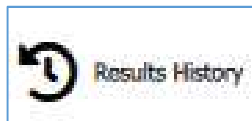


Рис. 9. Информационное окно



6.6. При выборе значка  вызывается содержимое файла “Results” (см. рис. 10).



Рис. 10. Содержимое файла “Results”




6.7. При выборе значка  открывается окно настроек (Рис. 11), позволяющее устанавливать множители dB и dC отдельно для женских и мужских голосов (см. формулу (2) в разделе 3).

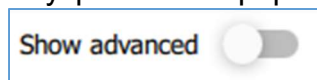


Рис. 11. Окно настроек - “Settings”

На Рис.11 по умолчанию даны значения этих множителей, выбранные разработчиками экспериментально.

ВНИМАНИЕ!

Внутренняя информация, вызываемая пометкой в правом верхнем углу



, предназначена только для разработчика.

Изменения её содержимого пользователем крайне нежелательно!