



**Программа для ЭВМ  
«RMS Acoustic Signature Detector (RMS-ASD)»**

**Описание функциональных характеристик программного  
обеспечения и информация, необходимая для установки и  
эксплуатации**

На 23 листах

2024

## АННОТАЦИЯ

В документе представлено описание функциональных характеристик программы для ЭВМ “RMS Acoustic Signature Detector (RMS-ASD)” (далее также – ПО, программа, программное средство). ПО предназначено для эмуляции работы программно-аппаратного комплекса обнаружения источника звука по акустическим сигнатурам и оценки направления на источник звука, характеристика которого задана исходным набором акустических сигнатур. Программа может использоваться в лабораторных стендах для исследования акустической обстановки на предмет идентификации источника звука с определенными акустическими параметрами и оценки его положения.

Функциональные возможности программы: загружает исходные акустические данные в различных форматах; определяет ключевые характеристики (сигнатуры) звукового сигнала; на основе обработанных оператором доверенных сигналов исследуемого источника звука формирует набор опорных сигнатур для последующего решения задачи обнаружения; по выбранной модели производит вычисление сигнатур произвольного звукового сигнала и выполняет сравнение полученных сигнатур с опорными, принимая решение о наличии/отсутствии искомого источника звука, оповещает об этом оператора; загружает исходные акустические данные в различных форматах, записи которых сформированы синхронно от трех микрофонов; определяет направление на источник звука в заданный момент времени; на заданном временном интервале анализа реализует процедуру фильтрации полученных оценок направлений на источник звука.

## СОДЕРЖАНИЕ

1 НАЗНАЧЕНИЕ ПРОГРАММНОГО СРЕДСТВА	4
2 УСЛОВИЯ ВЫПОЛНЕНИЯ ПРОГРАММНОГО СРЕДСТВА	5
3 ВЫПОЛНЕНИЕ ПРОГРАММНОГО СРЕДСТВА	5
3.1 Установка программного средства	5
3.2 Первоначальная настройка программного средства	6
3.3 Работа с программным средством	7
3.3.1 Загрузка исходных данных	7
3.3.2 Получение признаков	8
3.3.3 Обучение	13
3.3.4 Обнаружение источника	19
3.3.5 Определение направления на источник	21
3.3.6 Контрольный пример	20

# 1 НАЗНАЧЕНИЕ ПРОГРАММНОГО СРЕДСТВА

Программа предназначена для выделения в записанных акустических сигналах ключевых параметров (сигнатур) заданного источника звука и определения направления на источник звука на основе сравнения данных, записанных тремя разнесенными приемниками, оповещает об этом оператора. На основе обработанных оператором доверенных сигналов исследуемого источника звука формирует набор опорных сигнатур для последующего решения задачи обнаружения; по выбранной модели производит вычисление сигнатур произвольного звукового сигнала и выполняет сравнение полученных сигнатур с опорными по Байесовскому критерию, принимая решение о наличии/отсутствии искомого источника звука, оповещает об этом оператора.

ПО обеспечивает:

- загрузку исходных данных с записью акустических сигналов;
- определение ключевых параметров (сигнатур) источника акустического сигнала;
- обработку сигнатур с целью формирования опорной структуры, применяемой для принятия решения по наличию/отсутствию источника звука;
- обработку записи звукового сигнала с целью обнаружения источника звука по заданным сигнатурам;
- определение направления на источник звукового сигнала псевдодальномерным методом;
- фильтрацию полученных оценок направлений на источник звукового сигнала на временном интервале анализа.

## 2 УСЛОВИЯ ВЫПОЛНЕНИЯ ПРОГРАММНОГО СРЕДСТВА

Программа функционирует на вычислительных платформах x64 под управлением ОС семейства Linux (рекомендуется Debian 12, Ubuntu 22.04).

Минимальные системные требования:

CPU Intel Core i3 II поколения, RAM 4 ГБ, ROM 100 Мб.

Рекомендуемые системные требования:

CPU Intel Core i7 4 поколения, RAM 16 ГБ, ROM 2 Гб.

## 3 ВЫПОЛНЕНИЕ ПРОГРАММНОГО СРЕДСТВА

### 3.1 Установка программного средства

Для подготовки системного окружения для работы программы требуется выполнить следующие шаги (на примере ОС Ubuntu 22.04):

1. Установить ОС по умолчанию (на физическое аппаратное окружение или виртуальное (рекомендуется использовать виртуальные машины VirtualBox))
2. Установить .NET Core Runtime:

- `wget https://dot.net/v1/dotnet-install.sh -O dotnet-install.sh`
- `chmod +x ./dotnet-install.sh`
- `./dotnet-install.sh --channel 6.0`
- `echo 'export DOTNET_ROOT=$HOME/.dotnet' >> ~/.bashrc`
- `echo 'export PATH=$PATH:$DOTNET_ROOT:$DOTNET_ROOT/tools'`

`>> ~/.bashrc`

Примечание: после установки необходимо перезайти в систему.

0. Скачать архив с программой по адресу <http://rms-algo.ru/asd>

- 0. Распаковать архив с программой в домашнюю папку
- 0. Запуск программы:
  - ~\$ ./ASD\_1.0.0/Release/net6.0/ASD

### 3.2 Первоначальная настройка программного средства

ПО состоит из следующих логических модулей:

- модуль графического интерфейса;
- модуль сохранения и загрузки исходных данных и настроек;
- модуль расчета сигнатур;
- модуль компоновки сигнатур;
- модуль обнаружения;
- модуль определения направления;
- модуль журналирования.

Модуль графического интерфейса обеспечивает взаимоувязку остальных модулей, передачу данных между ними, а также сигналов управления от элементов интерфейса пользователя и отображения результатов расчета.

Модуль сохранения и загрузки позволяет сохранять рассчитанные данные, а также текущие настройки алгоритмов обработки между сеансами работы с ПО.

Модуль расчета сигнатур загружает файл с данными записанного сигнала, выделяет значимые сигнатуры и сохраняет их в оперативной памяти для последующей обработки модулем компоновки сигнатур.

Модуль компоновки сигнатур получает значимые сигнатуры от модуля расчета и компоует их в опорную структуру, на основе которой модуль обнаружения будет принимать решение о наличии источника сигнала.

Модуль обнаружения загружает значимые сигнатуры по конкретной записи и производит расчет, в результате которого принимает решение о наличии источника искомого сигнала в записи, используя опорную структуру, полученную

модулем компоновки.

Модуль определения направления загружает исходные данные от 3х источников (3 файла) и производит расчет, в результате которого принимает решение о направлении на источник звука. Также по набору полученных оценок модуль реализует процедуру фильтрации полученных оценок направлений на источник звукового сигнала на временном интервале анализа.

Модуль журналирования информирует пользователя о ходе расчетов и возникающих ошибках.

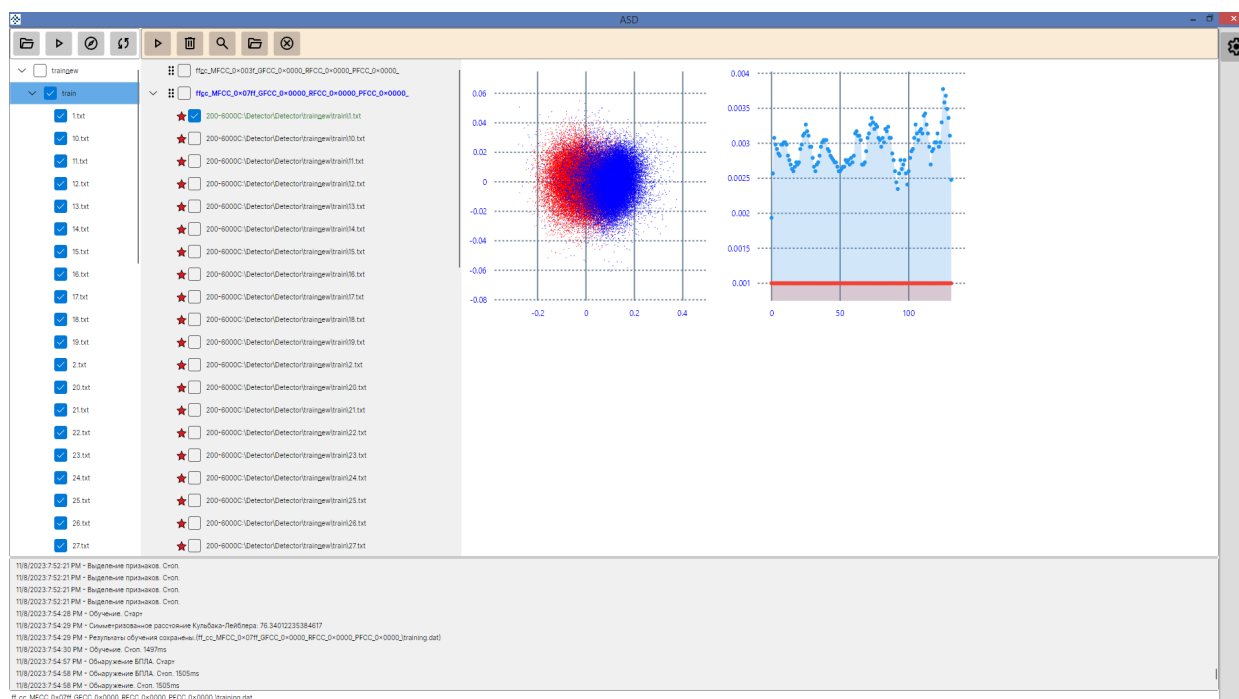



Рисунок 1. Внешний вид ПО

### 3.3 Работа с программным средством

#### 3.3.1 Загрузка исходных данных

Для начала работы оператору необходимо подключить папку с исходными данными (записанными акустическими сигналами).

Для этого следует нажать на кнопку «Открыть» (  ) (рисунок 2) и выбрать

путь к папке.

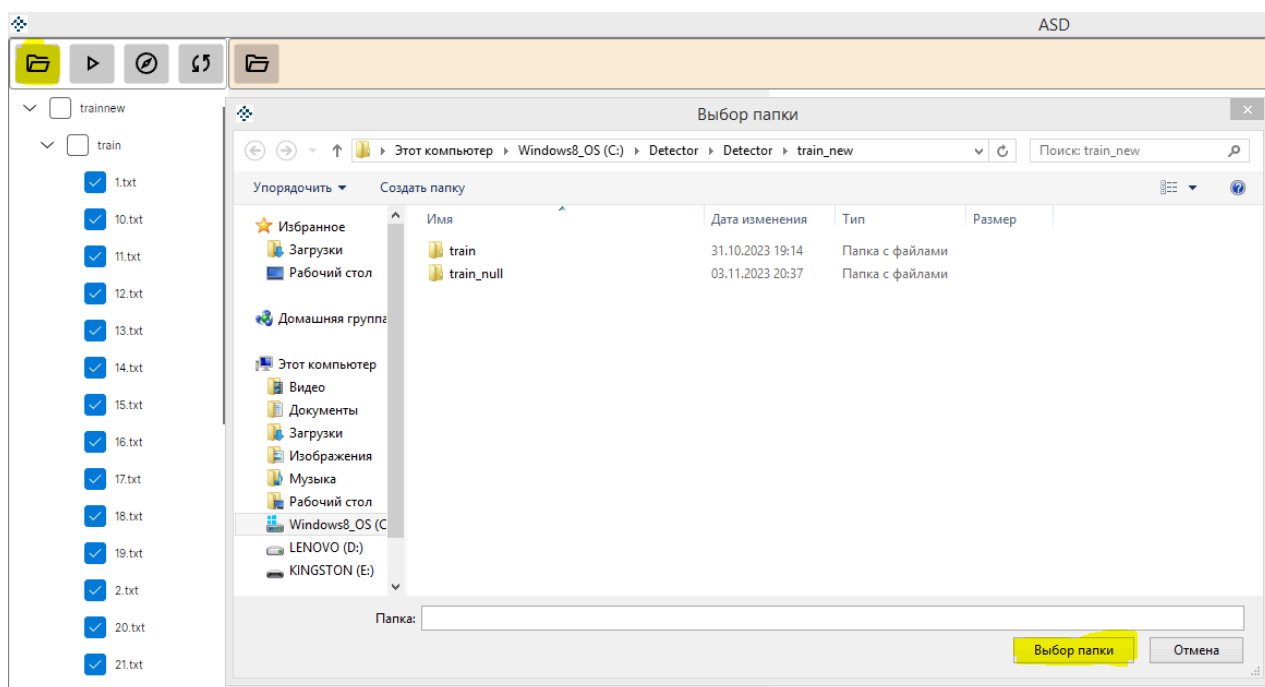



Рисунок 2. Подключение исходных данных

### 3.3.2 Получение признаков

Далее на правой вкладке (рисунок 3) необходимо нажать на кнопку «Настройка» (  ) и в раскрывшемся меню установить расчетные параметры на вкладке «Обработка» (рисунок 3), в соответствии с которыми будет произведено формирование доверительных сигнатур источника звука и последующая оценка его наличия/отсутствия.

Параметры включают в свой состав:

- размер фрейма в отсчетах звукового сигнала, на которых будет производится вычисление вектора ключевых параметров источника звука – сигнатур (по умолчанию задан значением 1024);
- процент перекрытия фреймов (по умолчанию задан значением 0,5);



– число коэффициентов полосового фильтра (ПФ), выделяющего полезную часть акустического сигнала в установленной полосе частот (по умолчанию задан значением 111) и границы нижней и верхней полосы частот ПФ в Гц полезной части акустического сигнала (по умолчанию заданы значениями 900 и 15000 соответственно). Пользователю предоставляется возможность отключения соответствующей функции фильтрации при снятии флаговой кнопки – «Включить фильтрацию»;

– число спектральных полос фильтров, составляющих координатный вектор сигнатур, применяемых при вычислении акустических параметров источника звука в мел-шкале, гамматон-шкале и барк-шкале (по умолчанию задан значением 12);

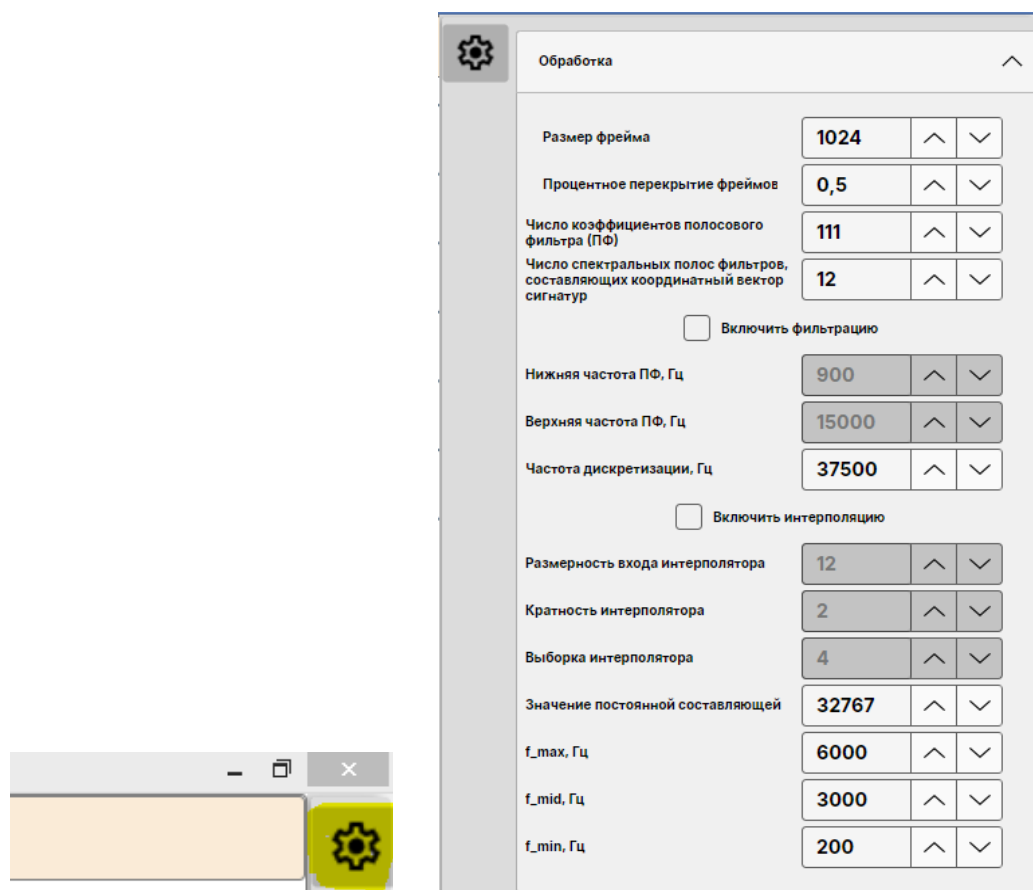


Рисунок 3. Установка параметров, по которым будет производиться обработка записанных звуковых файлов при выделении акустических сигнатур

– параметры передискретизации исходного сигнала, устанавливаемые значениями: размерности входа интерполятора, кратности интерполятора и

выборки интерполятора (при установленных на рисунке 3 значений сигнал будет передискретизирован при уменьшении частоты дискретизации в два раза). Пользователю предоставляется возможность отключения соответствующей функции передискретизации при снятии флаговой кнопки – «Включить интерполяцию»;

- значение постоянной составляющей, которая будет вычтена из исходного сигнала на этапе его предварительной обработке (по умолчанию задана – 32767)

- $f\_max$ ,  $f\_mid$  и  $f\_min$  в Гц – соответствующие значения верхней, центральной и нижней частот в области которых будет производиться вычисление координатного вектора сигнатур. При этом частоты  $f\_max$  и  $f\_min$  устанавливают полосу, в которой выбираются спектральные полосы фильтров в мел-шкале, гамматон-шкале и барк-шкале, а области  $f\_max$  и  $f\_mid$ ,  $f\_mid$  и  $f\_min$  задают участки, в спектральной области которых будет производиться вычисление перплексий высокочастотной, низкочастотной части спектра соответственно.


После установки параметров на вкладке «Обработка» следует уточнить значения координатного вектора сигнатур, которые будут составлять основные акустические параметры оценки источника звука при его последующей оценке наличия/отсутствия. Для этого в окне настроек необходимо открыть вкладку «Координатный вектор сигнатур» (рисунок 4) и выбрать параметры координатного вектора, которые включают в свой состав:

- перплексия высокочастотной части спектра;
- перплексия низкочастотной части спектра;
- MFCC – melfrequency cepstral coefficients (мел-частотные кепстральные коэффициенты);
- GFCC – gammatone frequency cepstral coefficients (гамматон-частотные кепстральные коэффициенты);
- RFCC – модифицированные синусоидальные барк-шкалированные частотные кепстральные коэффициенты;
- PFCC – перплекссионные барк-шкалированные частотные кепстральные коэффициенты.

Параметр	Значение
MFCC (0x0..7FFF)	0x7FF
GFCC (0x0..FFFF)	0x0
RFCC (0x0..FFFF)	0x0
PFCC (0x0..FFFF)	0x0

Рисунок 4. Установка элементов координатного вектора сигнатур

Включение соответствующих сигнатур в набор параметров оценки и обучения выполняется установкой флаговых кнопок и заданием включаемых элементов в векторах параметров в 16й системе исчисления (например, для заданного на рисунке 3 значения «0x3F» для MFCC будут выбраны первые шесть элементов – 111111, а для выбора восьмого и первых шести элементов 10111111 необходимо установить значение «0xBF»).

По завершению установки параметров обработки необходимо выбрать файлы, в которых требуется выделить сигнатуры. Соответствующий выбор реализуется установкой флаговых кнопок в дереве исходных данных (рисунок 5). После чего в группе «Исходные данные» необходимо нажать кнопку «Получение признаков» (  ).

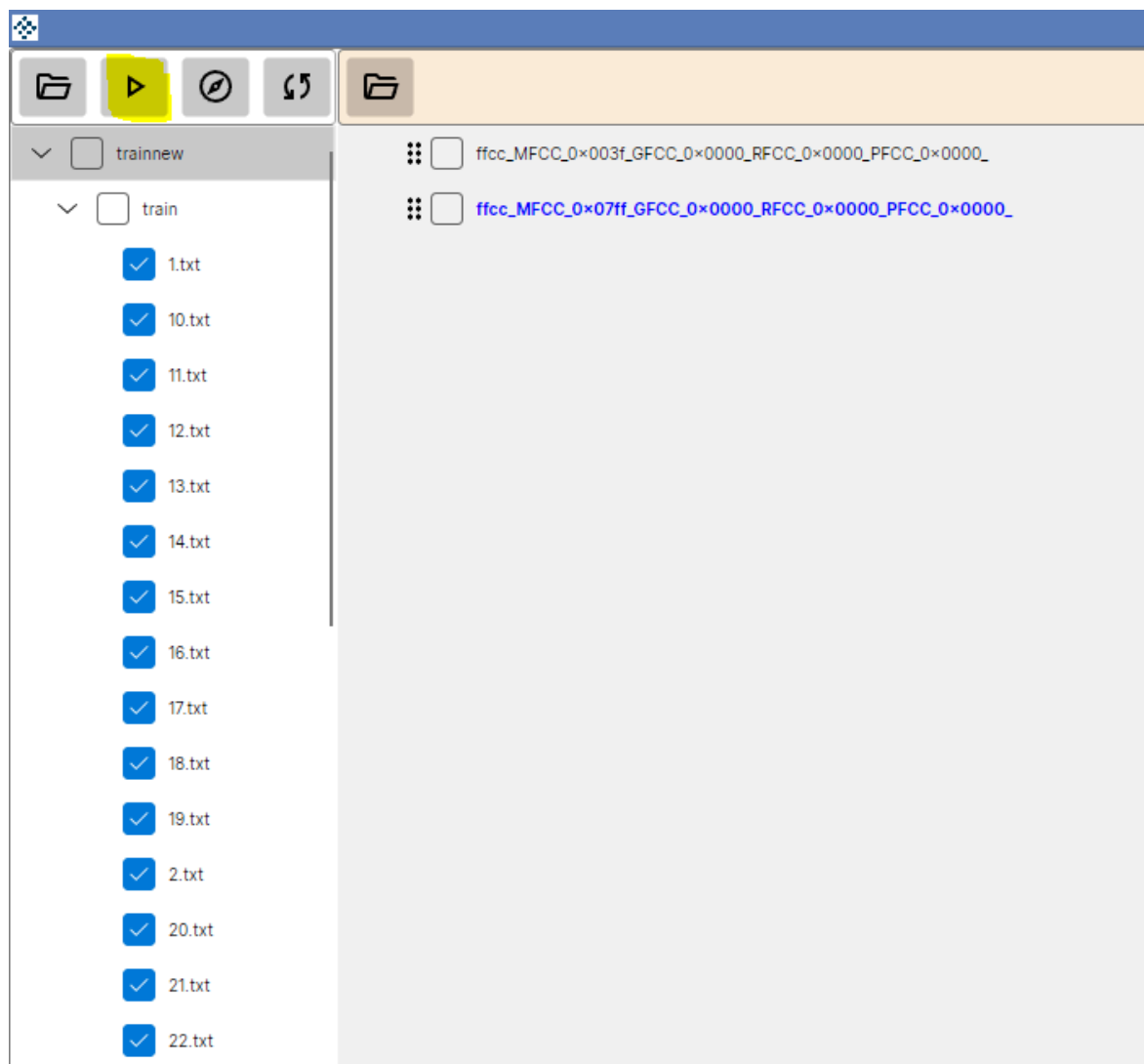


Рисунок 5. Запуск обработки с целью получения значимых признаков (сигнатур) сигналов

После завершения обработки данные с сигнатурами появятся в виде узлов в дереве «Признаки», а особенности проведенных расчетов уточнятся на вкладке «Лог» (рисунок 6).

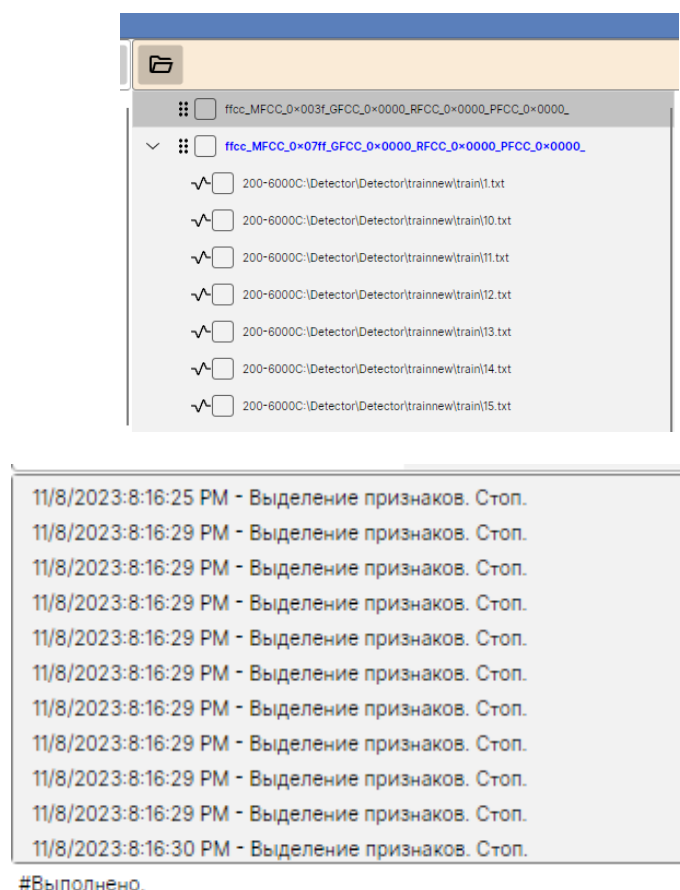






Рисунок 6. Сформированные признаки (сигнатуры) сигналов

### 3.3.3 Обучение

Полученные данные могут быть использованы для формирования опорной структуры сигнатур (если программа запущена впервые – этот шаг обязателен). Для этого необходимо выбрать признаки – эталоны, на основе которых алгоритм должен сформировать опорную структуру. Соответствующий выбор реализуется установкой флаговых кнопок в дереве «Признаки» (рисунок 7). Затем необходимо уточнить, какие признаки (обработанные файлы) содержат сигнатуры исследуемого источника звука (положительные), а какие включают ложные данные (отрицательные). Для определения положительных данных необходимо соответствующие отмеченные флаговыми кнопками файлы в дереве «Признаки» выбрать нажатием левой клавишей курсором мыши слева от названия и нажать на клавишу «PgUp». Определение отрицательных данных выполняется аналогичным

образом при нажатии на клавишу «PgDn». Правильное определение соответствующих положительных и отрицательных данных в программе идентифицируется красной (   200-6000C:\Detector\Detector\trainnew\train\6.txt ) и синий (   200-6000C:\Detector\Detector\trainnew\train\_null\1.txt ) подсветкой в значках рядом с названием соответствующих признаков (рисунок 7).

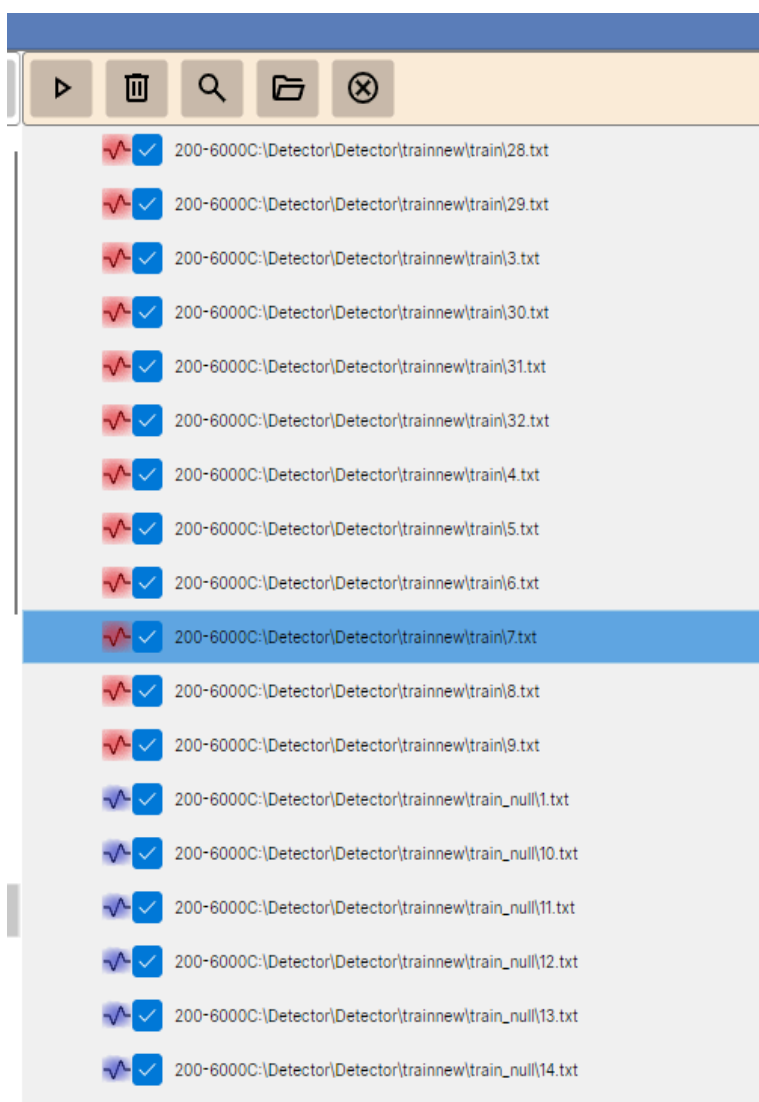


Рисунок 7. Старт формирования опорной структуры сигнатур

Поле завершения выбора файлов обучения в группе «Признаки» в окне

«Настройки» необходимо открыть вкладку «Оценка» (рисунок 8) и внести расчетные данные:

- размерность признакового пространства, в котором будет выполняться оценивания наличия/отсутствия источника звука по байесовскому правилу (по умолчанию значение равно 3)

- порог оценивания, устанавливающий минимальное значение вероятности наличия источника звука, вычисляемому по байесовскому правилу (по умолчанию задан 0,001);

- разрешение формируемой эмпирической функции плотности вероятности (ЭФПВ), которая формируется по рассчитанным наборам доверительных сигнатур источника звука за вычетом отрицательных данных (по умолчанию значение задано равным 120). Это значение определяет порядок аппроксимации прямоугольными базисными функциями эмпирическую функцию плотности эталонов. По соответствующей аппроксимации эмпирической функции плотности вероятности будет выполняться оценка наличия/отсутствия источника звука по критерию Байеса;

- минимальное число фреймов оценки  $x16$  – число фреймов, применяемых при оценке источника звука (минимальный размер окна оценки), устанавливающий минимальное значение набора векторов признаков, по которому выполняется однократная оценка наличия/отсутствия источника звука (по умолчанию задано 3);

- число кратности 16 точкам, определяющим максимальное число полуфреймов в наборе (максимальный размер окна оценки);

- максимальное число фреймов оценки  $x16$  – число фреймов, применяемых при оценке источника звука (максимальный размер окна оценки), устанавливающий максимальное значение набора векторов признаков, по которому выполняется однократная оценка наличия/отсутствия источника звука (по умолчанию задано 5).

Оценка	
Размерность признакового пространства	3
Порог принятия решения	0,001
Разрешение формируемой ЭФПВ	120
Минимальное число фреймов оценки x16	3
Число кратности 16 точкам, определяющее максимальное число полуфреймов в наборе (больше или равно 2м)	16
Максимальное число фреймов оценки x16	5

Рисунок 8. Ввод параметров оценивания источника звука

Затем следует выполнить ввод параметров обучения. Для этого в окне «Настройки» следует перейти на вкладку «Обучение» (рисунок 9).

Параметры обучения включают в свой состав:

- максимальное число итераций при вычислении собственных векторов ковариационной матрицы признаков, используемых при построении проекции в линейном дискриминантном анализе – LDA (по умолчанию значение – 1000);
- данных уточняющих градиентных алгоритмов, которые переопределяют проецирующие вектора LDA по критериям: 1 этап оптимизации максимизирует расстояние Кульбака–Лейблера между вероятностями полезных и отрицательных данных, полагая то, что они распределены по многомерному нормальному закону с вычисляемыми моментами (векторы математического ожидания и ковариационными матрицами); 2 этап оптимизации максимизирует расстояние Кульбака–Лейблера между эмпирическими функциями максимального правдоподобия конструируемыми нормальным распределением или распределением Стюдента. Исключение оптимизационных алгоритмов выполняется снятием соответствующей флаговой кнопки «Использовать уточняющий градиентный алгоритм». Данные остановки градиентных алгоритмов устанавливаются максимальным числом итераций, погрешностью вычисления оптимума. Скорость обучения задается начальным шагом градиентной оптимизации и сглаживающим параметром, которые соответствуют входным



параметрам оптимизации метода Нестерова.

The screenshot shows a software window titled "Обучение" (Training) with a collapse icon. It contains several sections for configuring training parameters:

- Максимальное число итераций при определении собственных чисел и векторов в методе LDA:** A numeric input field set to 1000 with up and down arrow buttons.
- Использовать уточняющий градиентный алгоритм:** An unchecked checkbox.
- 1 Этап оптимизации:** A collapsible section containing:
  - Количество итераций (Ng):** 2000
  - Погрешность вычисления оптимума:** 1E-15
  - Погрешность вычисления длины градиентного шага:** 1E-08
  - Скорость обучения:** 0.2
  - Сглаживающий параметр:** 0.85
- Уточняющий алгоритм:** A collapsible section containing:
  - Количество итераций (Ng):** 10
  - Погрешность вычисления оптимума:** 1E-15
- 2 Этап оптимизации:** A collapsible section containing:
  - Критерий:** Two radio buttons. "По критерию Гаусса" is unselected, and "По критерию Стюдента" is selected.
  - Количество итераций (Ng):** 10000
  - Погрешность вычисления оптимума:** 1E-15
  - Погрешность вычисления длины градиентного шага:** 1E-08
  - Скорость обучения:** 0.04
  - Сглаживающий параметр:** 0.85
  - Число степеней свободы для критерия Стюдента:** 2

Рисунок 9. Ввод параметров обучения

После установки параметров обучения необходимо произвести расчет обучающих сигнатур нажатием на кнопку обучение (▶). В результате в папке с исполняемым файлом программы в папке, название которой соответствует выбранной группе сигнатур (для рассматриваемого на рисунке 6 случая папка «ff\_cc\_MFCC\_0x07ff\_GFCC\_0x0000\_RFCC\_0x0000\_PFCC\_0x0000\_»), сформируется файл training.dat, который будет использоваться для решения

задачи обнаружения до следующего запуска процесса обучения. Информация о соответствующей операции будет отображения в окне «Лог» (рисунок 10).

```
11/8/2023 8:49:51 PM - Симметризованное расстояние Кульбака-Лейблера: 75.8038094632008  
11/8/2023 8:49:51 PM - Результаты обучения сохранены (ft_cc_MFCC_0x07ff_GFCC_0x0000_RFCC_0x0000_PFCO_0x0000_training.dat)  
11/8/2023 8:49:51 PM - Обучение. С топ. 1661ms  
ft_cc_MFCC_0x07ff_GFCC_0x0000_RFCC_0x0000_PFCO_0x0000_training.dat
```

Рисунок 10. Сведения о завершении процедуры расчета обучающих данных в окне «Лог»

Также по завершению расчета эталонов в группе объектов «Диаграммы» будут отражения графики точек проекций вычисленных сигнатур обучающих звуковых сигналов (рисунок 11).

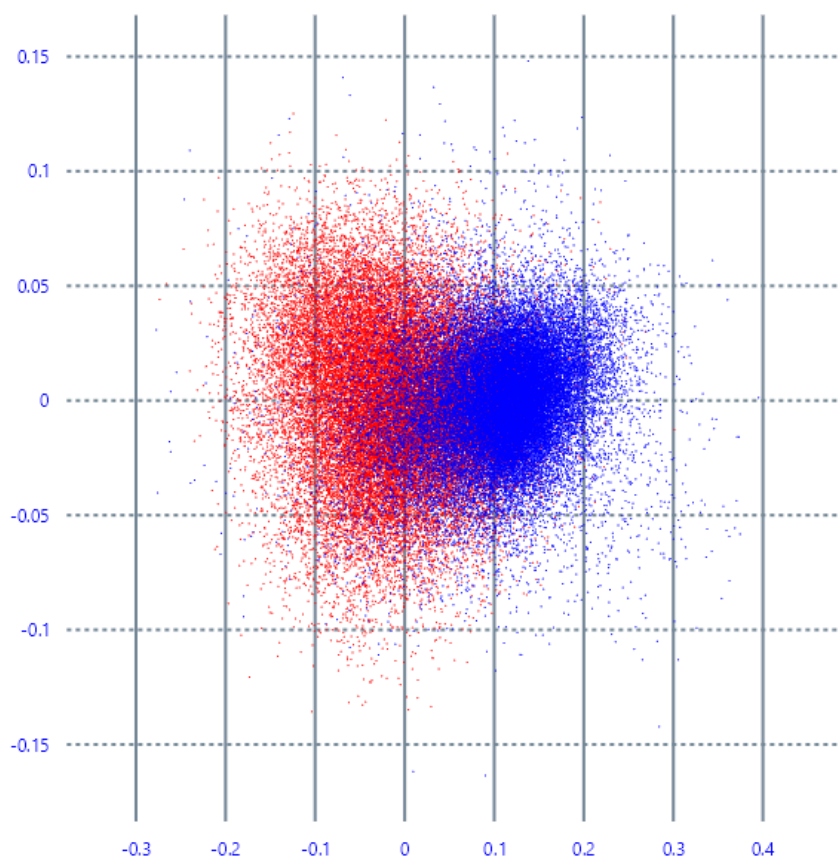


Рисунок 11. Результат расчета точек проекции эталонов, по которым

осуществляется аппроксимаций эмпирической функции плотности вероятности

### 3.3.4 Обнаружение источника звука

Для обнаружения источника звука, обладающего заданными характеристиками (сигнатурами), также используются данные в окне «Признаки» (рисунок 12), которые необходимо вычислить нажатием на кнопку «Получение признаков». Для корректного вычисления набор признаков (сигнатур) должен соответствовать всем параметров, задаваемым на этапе получения эталонов.

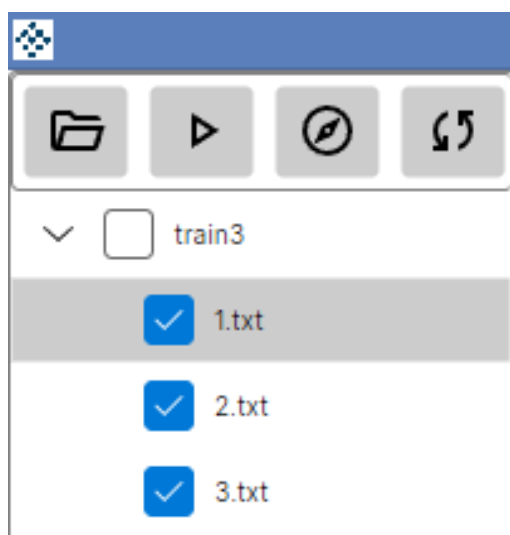


Рисунок 12. Запуск процесса расчета признаков в звуковых сигналах

После расчета признаков оператор должен выбрать один или несколько узлов в дереве «Признаки» (рисунок 13) и нажать кнопку «Поиск сигнатур»



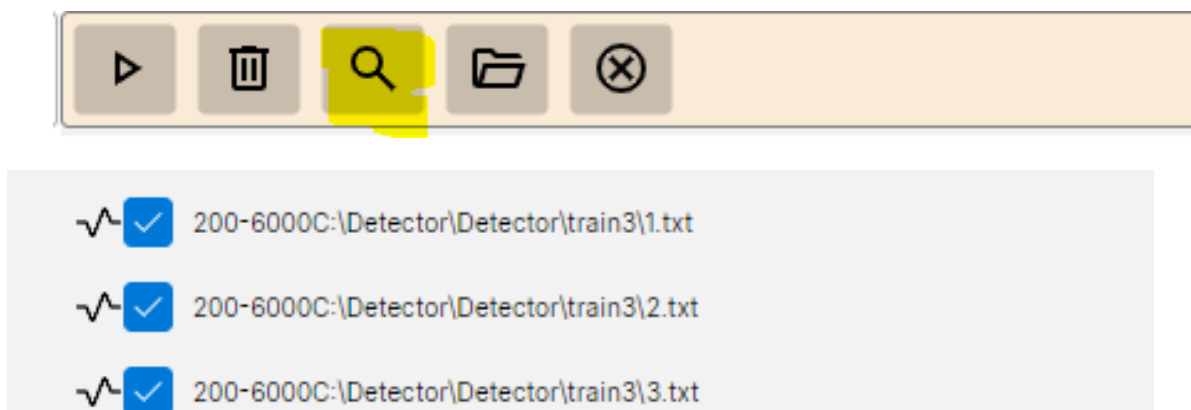


Рисунок 13. Запуск процесса поиска источника звука

В результате работы алгоритма пользователь получит диаграммы, которые отражают вероятностную характеристику присутствия источника звука по времени (рисунок 14).

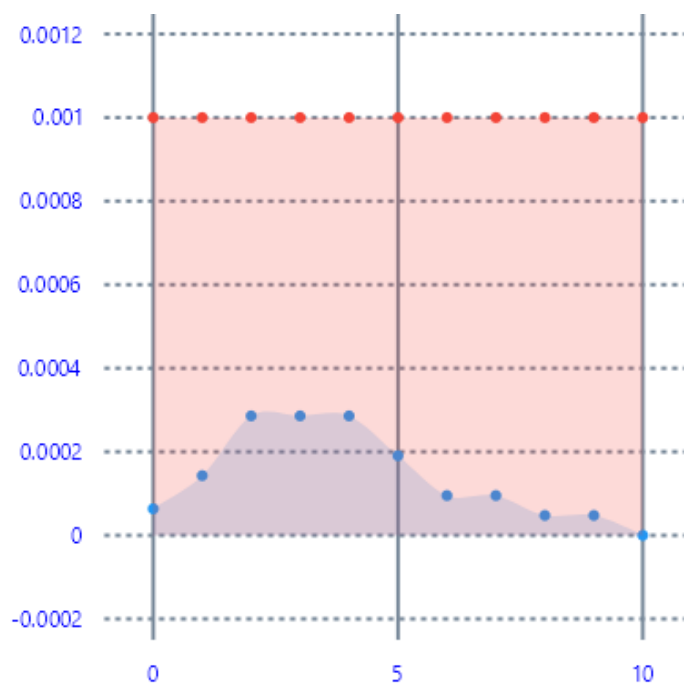


Рисунок 14. Итоговая диаграмма

### 3.3.5 Определение направления на источник

Для определения направления оператору необходимо на вкладке «Обработка» (рисунок 3) установить расчетные параметры, в соответствии с которыми будет произведена предварительная обработка звуковых файлов, предполагающая заданий:

- число коэффициентов фильтра, выделяющего полезную часть акустического сигнала в установленной полосе частот (по умолчанию задан значением 111) и границы нижней и верхней полосы частот в Гц полезной части акустического сигнала (по умолчанию заданы значениями 900 и 15000 соответственно). Пользователю предоставляется возможность отключения соответствующей функции фильтрации при снятии флаговой кнопки – «Включить фильтрацию»;

- число фильтров, применяемых при вычислении акустических сигнатур источника звука в мел-шкале, гамматон-шкале и барк-шкале – JJ (по умолчанию задан значением 12);

- параметры передискретизации исходного сигнала, устанавливаемые значениями: размерности входа интерполятора, кратности интерполятора и выборки интерполятора (при установленных на рисунке 3 значений сигнал будет передискретизирован при уменьшении частоты дискретизации в два раза). Пользователю предоставляется возможность отключения соответствующей функции передискретизации при снятии флаговой кнопки – «Включить интерполяцию»;

- значение постоянной составляющей, которая будет вычтена из исходного сигнала на этапе его предварительной обработке (по умолчанию задана – 32767).


По завершению установки параметров обработки необходимо перейти на вкладку «Направление» (рисунок 15) и задать параметры, в соответствии с которыми будет выполняться процедуры оценки и фильтрации оценок

направлений на источник звука.

Параметр	Значение	Вверх	Вниз
Скорость звука м/с	331,46	^	v
Исходная частота дискретизации, Гц	37500	^	v
Смещение в отсчетах	8	^	v
Число интервалов	30	^	v
Используемое число интервалов	25	^	v
Alpha	0,46	^	v
Число коэффициентов фильтра	64	^	v
Число отсчетов влево-вправо	4	^	v
Число элементов интерполяции	5	^	v
Диаметр описанной окружности	0,075	^	v
Расстояние, м	100	^	v
Число итераций	50	^	v
Размер окна в фильтре Винера	32	^	v

Сохранить

Рисунок 15. Установка параметров, по которым будет производиться оценка и фильтрация оценок направлений на источник звука

После установки расчетных параметров следует выбрать файлы (3 файла), в которых содержится синхронная запись одного сигнала тремя соответствующими приемниками (микрофонами). Данный выбор реализуется установкой флаговых кнопок в дереве исходных данных (рисунок 16). После чего в группе «Исходные данные» необходимо нажать кнопку «Определить направление» (  ).

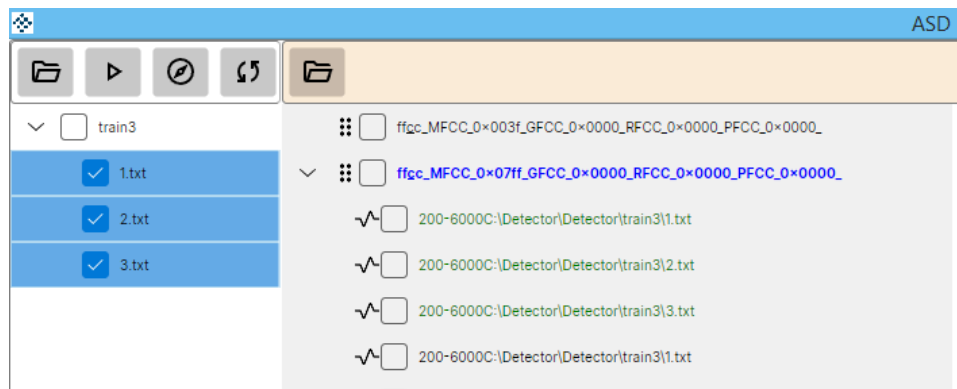


Рисунок 16. Установка параметров, по которым будет производиться оценка и фильтрация оценок направлений на источник звука

В результате работы алгоритма пользователю отобразиться форма, отражающая диаграммы, которые представляют направления (углы в сферической системе координат – отклонения и обхода) на источник звука по времени (рисунок 17).



Рисунок 17. Результат работы программы по оценке направлений на источник звука